

Run on: December 19, 2002, 02:23:32 ; Search time 2737 Seconds

14758.748 Million cell updates/sec

Title: US-09-905-674-1

Sequence: 1 cttcctgcgcgagccggc.....gggaqcqnaacqttgccc 1388

Scoring table: IDENTITY\_NUC

Searched: 2054640 seqs, 14551402878 residues

total number of hits satisfying chosen parameters: 4109280

Maximum DB seq length: 20000000000

Post-processing:	Minimum Match	0%
------------------	---------------	----

Listing first 45 summaries

Database

1:	gb_da:*
2:	gb_htg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pro:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vl:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*
29:	em_vl:*
30:	em_htg_hum:*
31:	em_htg_in:*
32:	em_htg_other:*
33:	em_htg_mus:*
34:	em_htg_pln:*
35:	em_htg_rod:*
36:	em_htg_mam:*
37:	em_htg_vrt:*
38:	em_sy:*
39:	em_higo_hum:*
40:	em_higo_mus:*
41:	em_higo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1385.2	99.8	1388	6	AX440923	AX440923 Sequence
2	1352.2	96.0	2556	9	HSN801611	AL136638 Homo sapi
3	1323.4	95.3	2553	9	AFJ11903	AFJ11903 Homo sapi
4	1214	87.5	2502	9	BC002920	BC002920 Homo sapi
5	816.2	58.8	2428	10	BC025568	BC025568 Homo sapi
6	815.2	58.7	2498	10	BC026161	BC026161 Mus muscu
7	811.8	58.5	2500	10	BC026574	BC026574 Mus muscu
8	660	47.6	2605	9	AK055330	AK055330 Homo sapi
9	541.8	39.0	161804	9	AC021028	AC021028 Homo sapi
10	362	26.1	813	6	AX420466	AX420466 Sequence
11	362	26.1	3184	6	AX420466	AX420466 Sequence
12	348.8	25.1	1516	10	BC010346	BC010346 Mus muscu
13	345.2	24.9	2426	10	BC010405	BC010405 Homo sapi
14	340.2	24.5	2175	10	AF121344	AF121344 Mus muscu
15	337.4	24.3	1405	9	AF065389	AF065389 Homo sapi
16	337.4	24.3	1416	9	BC009704	BC009704 Homo sapi
17	303	21.8	1408	9	AF053455	AF053455 Homo sapi
18	230	16.6	864	6	AX434015	AX434015 Sequence
19	216.4	15.6	1567	6	AX061778	AX061778 Sequence
20	206.4	14.9	557	6	AX079573	AX079573 Sequence
21	203	14.6	4445	6	AX024427	AX024427 Homo sapi
22	178	12.8	852	6	AX247836	AX247836 Sequence
23	177.6	12.8	1995	10	BC024685	BC024685 Mus muscu
24	158.4	11.4	179772	2	AC108430	AC108430 Mus muscu
25	135.2	9.7	143717	2	AC123389	AC123389 Rattus no
26	134.8	9.7	143717	2	AC123389	AC123389 Rattus no
27	134.8	9.7	167399	3	AC103212	AC103212 Rattus no
28	130.4	9.4	1216	3	AF274013	AF274013 Drosophill
29	124.4	9.0	1703	9	AF120266	AF120266 Homo sapi
30	124.4	9.0	1726	9	BC003157	BC003157 Homo sapi
31	124.4	9.0	1734	9	BC004161	BC004161 Homo sapi
32	122.8	8.8	1778	6	AX376210	AX376210 Sequence
33	122.4	8.8	1661	10	BC003872	BC003872 Mus muscu
34	113.6	8.2	1601	9	AF325213	AF325213 Homo sapi
35	113.4	8.2	368	9	AF174603	AF174603 Homo sapi
36	111.6	8.1	1693	9	BC032802	BC032802 Homo sapi
37	111.6	8.0	1646	10	BC019991	BC019991 Mus muscu
38	105	7.6	167399	2	AC103212	AC103212 Rattus no
39	95.4	6.9	394	6	AX072723	AX072723 Sequence
40	91.8	6.6	389	6	AX072711	AX072711 Sequence
41	91.8	6.6	102468	2	AC009984	AC009984 Drosophill
42	91.8	6.6	106584	2	AC013977	AC013977 Drosophill
43	91.8	6.6	140973	3	AC008140	AC008140 Drosophill
44	91.8	6.6	159007	3	AC009219	AC009219 Drosophill
45	91.8	6.6	224795	3	AE003688	AE003688 Drosophill

## ALIGNMENTS

RESULT 1	AX440923			
LOCUS	AX440923	1388 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0206340.			
ACCESSION	AX440923			
VERSION	AX440923.1	GI:21665563		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1			
AUTHORS	Reinhardt, C.J. and Garcia, P.D.			
TITLE	Retraspan protein and uses thereof			
JOURNAL	Patent: WO 0206340-A 1 24-JAN-2002;			











Tissue Procurement: Jeffrey Green M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: nisc\_mgenhprt.nih.gov/  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Grantle, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
 Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Strantrop, S., Thomas, P.J., Touchman, J.W.,  
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: IRAX Plate: 56 Row: j Column: 21  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Genomescan gene  
 prediction.

# FEATURES

## source

Location/Qualifiers  
 1..2428  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="FVB/N"  
 /clone="MGC:36595 IMAGE:5322531"  
 /tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating  
 ductal carcinoma. 5 month old virgin mouse."  
 /clone\_lib="NCI CGAP\_Mam6"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 64..876  
 /codon\_start=1  
 /product="Similar to region containing hypothetical  
 protein MGC11352; slingshot"  
 /protein\_id="AAH25568.1"  
 /db\_xref="GI:19343884"  
 /translation="MAYRYSNAENSCWKYLLFESNYTFMLAGVFFLGVGMASEK  
 GVLSDLTAKYRLHIDPVVLAIVMGVMTLGFAGCVGLRENICLLKFFGCAIVLIF  
 FLEIAVAVALFLEODWVRDRFREFESNISKSYRDIIDLNLIDSLQKNAOCGAAYPE  
 DMDLNIVFNCSGASYSREKGVSCVDPKQKVVNTCCGYDVILOLKSMDKEIFLT  
 KGCIOALEMPLPNRNIYIVAGVFIAISLIDFGLFARLISIDIEAVKAGHNP"

## CDS

BASE COUNT 517 a 640 c 624 g 647 t  
 ORIGIN

Query Match 58.8%; Score 816.2; DB 10; Length 2428;  
 Best Local Similarity 83.2%; Pred. No. 9.5e-177;  
 Matches 1017; Conservative 0; Mismatches 188; Indels 18; Gaps 7;

QY 4 CCTGGCCGAGCCGGCGCGCCGCTGCGCCGCGCGCGGAGATTGCTTCTCGAC 63  
 DB 1 CCACGCGTCCGCGCGCGTCTGCTCCGCGCTGCGCGCGCGGCTCTCTTCGAC 60  
 QY 64 AAGATGCACTTTATAGTACTTAAGCCCAAGTCACTGCTGCTGCTCAAGTACTCTT 123  
 DB 61 AAGATGCACTTTATAGTACTTAAGCCCAAGTCACTGCTGCTGCTCAAGTACTCTC 120  
 QY 124 TTCAGTACAAATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183  
 DB 121 TTCAGTACAAATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 184 GCATGGAGCGAAAGGGTCTCTCGACCTACCAAGTACCGCGGATGCAATGATC 243  
 DB 181 GCATGGAGCGAAAGGGTCTCTCGACCTACCAAGTACCGCGGATGCAATGATC 240  
 QY 244 GACCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303  
 DB 241 GACCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 304 TGGTGGGGGCTTGGCGGAGATATCTGCTCAACTTTTCTGCTGCGACCATCGTG 363

DB 301 TGTGCGGGGCGCTCCGAGAGAACATCTGCTGCTCAAGTTTTTCTGTGGGCGATGTG 360  
 QY 364 CTCATCTTCTTCTCGAGACTGCTGTGGCGTGTGGCCTTCCGTTCCAGAGCTGGGTG 423  
 DB 361 CTCATCTTCTTCTCGAGACTGCTGTGGCGTGTGGCCTTCTTATTCAGAGCTGGGTG 420  
 QY 424 AGGAGCGGTTCCGGAGTCTTCTGAGAGCAATCAAGTCTACCGGAGATATGAT 483  
 DB 421 AGAGACCGGTTCCGGAGTCTTCTGAGAGCAATCAAGTCTACCGGAGATATGAT 480  
 QY 484 CTGCAAAACCTCATGACTCCCTTCAGAAAGCTAACCAAGTCTGTGGCGATATGGCCT 543  
 DB 481 CTGCAAAACCTCATGACTCCCTTCAGAAAGCTAACCAAGTCTGTGGCGATATGGCCT 540  
 QY 544 GAAGACTGGAGCCTCAAGCTCTACTTCAATTGCAAGCGGTGCACCTACAGCCGAGAAAG 603  
 DB 541 GAAGACTGGAGCCTCAAGCTCTACTTCAATTGCAAGCGGTGCACCTACAGCCGAGAAAG 600  
 QY 604 TGGGGGTGCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663  
 DB 601 TGTGGGTGCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
 QY 664 TGTGATATGATGTGAGATTACGCTGAGAGCAAGTGGATGAGTCCATCTTACGCAAA 723  
 DB 661 TGTGATATGATGTGAGATTACGCTGAGAGCAAGTGGATGAGTCCATCTTACGCAAA 720  
 QY 724 GGCTGATCCAGCGGCTGGAAGGCTGCTCCGCGGAGACATTTACATTTGGTGGGCTG 783  
 DB 721 GGATCATCCAGCGCTCTGGAAGGCTGCTGCGCAGAGAACATCTGATGTGGCGGTGC 780  
 QY 784 TTTCATGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843  
 DB 781 TTTCATGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
 QY 844 GACATCGAGGCACTGAAGCGCGCCATCATCTTCTGAGAGAGAGTGGAGAGCGCGAGC 903  
 DB 841 GACATCGAGGCACTGAAGCGCGCCATCATCTTCTGAGAGAGAGTGGAGAGCGCGAGC 899  
 QY 904 TGAGCCAGCTGGAGCGCAGAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963  
 DB 900 TGAGCCAGCTGGAGCGCAGAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 953  
 QY 964 GAGCGGAGCGCGCAGAGCGCAGAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023  
 DB 954 GACATTAGCGCACCAACCAAGCCAGTCTTCCGCAATGATGATATGATGATGATGAT 1013  
 QY 1024 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1083  
 DB 1014 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1070  
 QY 1084 TGCATCCCT--CTGAGTCTACCCAGAGACAGAGATGCTTATGATGAGAGTGTGA 1141  
 DB 1071 GGCATCTCTGAAGAGGCTTGGCCAGAGACAGAGATGTTTATATG--GGAGTAGTAA 1129  
 QY 1142 CTCTGAAGACAGAGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1201  
 DB 1130 CTCTGA---GGAAGATGCTCTCCATGCTGCTGAGAGAAAGCTTGA--TGGCCCTTATGA 1184  
 QY 1202 GCTCAAGCATGCTGAGGAGAC 1224  
 DB 1185 GCCCAAGACATCTGAGGCTTC 1207

RESULT 6  
 BC024611 2498 bp mRNA linear ROD 07-AUG-2002  
 LOCUS  
 DEFINITION  
 Mus musculus, Similar to region containing hypothetical protein  
 MGC11352; slingshot, clone MGC:28503 IMAGE:4188261, mRNA, complete  
 cds.  
 ACCESSION BC024611  
 VERSION BC024611.1 GI:1934053  
 KEYWORDS MGC.

SOURCE	ORGANISM
house mouse.	<i>Mus musculus</i>
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	
1 (bases 1 to 2458)	
Strausberg, R.	
Direct Submission	
Submitted (01-MAR-2002) National Institutes of Health, Mammalian	







```

/db_xref="GI:21524615"
/translation="MPGKHQHFQDEPEVCCGKRYFLGFENVFVWLGALFLAIGLWANG
EKGVLSNLSALDGLDIPVWLFVYGVMSYLAGAGCAGALRENTPLKFRSVFLG
IFLELATGILAFVFKDWIRDOJLNFINNVAAYRDDLOJLIDFADBYMSCCGARG
PNDMNLIYFNCDLNPBSRERGVPSCVRPAEDVLTGCGYIRUKLELEDDGSI
HTGCVQGFQFKWLODNLIVAGLVGALVIALOIFGLICLQNLVSDIKAVKANM"
BASE COUNT      675 a      883 c      945 g      681 t
ORIGIN
Query Match      26.1%; Score 362; DB 6; Length 3184;
Best Local Similarity 66.9%; Pred. No. 1.6e-72;
Matches 515; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 95 AAGTCAGCTGCTGTACAGTACTCTTTTACAGTCAACAATCATCTTGTGTTGGCTG 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 AGGTGGCTGTGGGAAATACCTGTTTGGCTTCAACATGTTCTGTGGGCTCTG 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 GAGTTGTTCTTGGATCGGGCTGTGGCATGAGCGAAAGAGGTGCTGTCCGACC 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 GAGCCCTTCTGCTGCTATCGGCTCTGGGCTGTGGGTGAGMAGGGCTTCTCGAACA 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 215 TCACCAAAAGTACCCGGATGCAATCGACCTGTGTGTCTGTCTGATGTTGGCG 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 TCTCAGCCCTGACAGATCTGGAGAGCCTTGACCCCTGTGGCTGTTGTGTGTTGAG 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 275 TGGTATGTTGACCCCTGGGTTGCCCGCTGCGGAGGCTGTGGGGGGAATATCGCT 334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 GCGTATCTGCGTGTGGCTGTTGCTGTGCTGATGAGGCGCTTCGGGAGAACCTTCC 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 335 TGCTCAACTTTTCTGTGGACACATGCTGCTCATCTTCTTCTGAGACCTGCTGAGCC 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 TGCTCAACTTTTCTGCTGTTCCGTGTTCTGCTCATCTTCTTCTGAGACTGGCAACAGGA 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 395 TGCTGGCCCTTCTGTTCCAGAGTGGGAGGAGGCGGCTCCGGAGTCTTTCAGAGCA 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 TCTTGGCCCTTGTCTTCAAGAGTGTGATGAGACACAGCTCAACCTTCTTATCAACACA 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 455 ACATCAAGTCTTACCGGAGCATATGATGCAACAAACCTATCGACTCCCTTCAGAAAG 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 AGCTCAAGGCTTACCGGAGCATATGACCTCCAGAACCTATGATGATTTGTCTCAGGAAT 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 515 CTAACAGTGTGTGGGAGCATATGGCCCTGAAGACTGGAGCTCAAGCTTCAATT 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 ACTGTCTTGTGTGAGACCCGAGGCCCAATGAGTGAATATATCTTACTTCAACT 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 575 GCAGCGTGGCAGCTACAGCCGAGAGAGTGCAGGAGTCCCTTCTCTGCTGCTGCGAG 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 GCAGTGAATTAACCCCGAGCGGAGCGCTGCGGGGTGCCCTTCTCTGCTGCTGAGG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 635 ATCTGCGCAAAAAGTGTGAACACAGTGTGATATGATGTAGAGATTCACTGAAGA 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 742 ACCCTGCGGAGATGTCTCTCAACACCGAGTGTGCTACGAGCTCCGCTCAAACTGAGC 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 695 GCAAGTGGGATGAGTCACTCTTCAAGAAAGGCTGATCCAGGCGCTGGAAAGCGGTCC 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 802 TGGAGCAGACAGGAGCTTCAATCCACACCAAGAGCTGCTGGGCAAGTTGAAGAGTGGCTG 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 755 CGCGAACAATTTTACATTTGTGCTGTGCGCTCTTCAATCGCATCTGCTGTGAGATATTG 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 862 AGGACAACCTGATTTGTGTGGCGGAGCTTCAATGGGATCGCCCTCTCCAGATCTTTG 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 815 GCATCTCTGCGAAGAGCGAGTCAATGATGAGAGGAGGAGGAGAGGCC 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 922 GCATCTGCTGGCCAGAACCTCGTGAATGATCAAGGAGTGAAGGCC 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
REMARK
COMMENT
MGC.
house mouse.
Mus musculus
Mus musculus
Mammalia: Eutheria: Rodentia: Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1516)
Strausberg, R.
Direct Submission
Submitted (05-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project url: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: g4pds-f@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-MGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://Image.lnl.gov
Series: IRAP Plate: 5 Row: f Column: 18.
location/Qualifiers
1. 1516
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="C57BL/6J"
/clone="MGC:6941 IMAGE:2811935"
/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
old, gross tissue."
/clone_11b="NCI-CGAP_Mam5"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
136..948
/codon_start=1
/product="RIKEN CDNA 2210021G21 gene"
/protein_id="AAH10346.1"
/db_xref="GI:16307593"
/db_xref="LocusID:74257"
/translation="MPGKHQHFQDEPEVCCGKRYFLGFENVFVWLGALFLAIGLWANG
EKGVLSNLSALDGLDIPVWLFVYGVMSYLAGAGCAGALRENTPLKFRSVFLG
IFLELATGILAFVFKDWIRDOJLNFINNVAAYRDDLOJLIDFADBYMSCCGARG
PNDMNLIYFNCDLNPBSRERGVPSCVRPAEDVLTGCGYIRUKLELEDDGSI
YTGCVQGFQFKWLODNLIVAGLVGALVIALOIFGLICLQNLVSDIKAVKANM"
BASE COUNT      277 a      417 c      461 g      361 t
ORIGIN
Query Match      25.1%; Score 348.8; DB 10; Length 1516;
Best Local Similarity 65.6%; Pred. No. 1.7e-69;
Matches 509; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 89 ACGCAAGTACAGTGTGAGTACAGTACCTCTTTCACATCAACATCATCTCTGAT 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 ACCCTGAGAGCTGCGTGTGGGAAATACCTTCTGTTGCTTCAACATGCTTCTGG 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 149 TGGCTGAGTGTCTTCTTGGAGTGTGGGCTGTGGGAGGAGGAGGAGGAGTGTCTG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 TGGCTGAGAGCTTCTTCTGCGATGCGCTCTGCGGAGGAGGAGGAGGAGGAGTCTG 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 209 CCGACCTCAACCAAGTACCCGAGTGCATGAGATCGACCTGTGAGTGTGCTGATG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 CCAATATCTGCGGCTGACATCTGCGGAGCTGTGACCCGCTGTGCTGTGTG 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 269 TGGCGGTGATGTTTCAACCTGGGAGTTCGCGGCTGTGGGAGGAGTGTGCGGAGAAAT 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 12

BC010346

LOCUS

DEFINITION

ACCESSION

VERSION

BC010346

IMAGE:2811935, mRNA, complete cds.

BC010346

BC010346.1

1516 bp

mRNA

linear

ROD 07-AUG-2002

Mus musculus, RIKEN CDNA 2210021G21 gene, clone MGC:6941

complete cds.

GI:16307592



Db	344	TGGGGGAGTCATGTCAGTGTGGGCTTCGGCCGGCTGCATTTGGGGCCCTTCGGGGAAAAA	403
QY	329	TCGCTGTGCATCACTTTTCTGTGGACACATCGTGTGCATCTTCTCTGGAGCTGGCTG	388
Db	404	CTTTCCTGCTCAAGTCTTCTGTGTGTCTCTCGGCGCCATCTCTTCTTCTGAGCTGGCG	463
QY	389	TGGCCGTGTGGCTTCTCTGTCCAGACATGGGTGAAGGACCGGTTCCGGAGATCTCTCG	448
Db	464	CGGGGATCTCGGCGCTTGTCTCTCAAGATTGGATTCGAGACCAACTTAATCTCTTATCA	523
QY	449	AGACCAATCAATCTCTACCGGAGCATATGCATGTGCAAAAACCTATCGACTTCCCTT	508
Db	524	ACAACATGTCAAGGCTTACCGGATGATCTCGACCTTCAAGACCTATGAGCTTGGCTC	583
QY	509	AGAAAGCTAACCACTGCTGTGGCGCATATGGCCCTTGAAGACTGGGACCTTAACGTCTACT	568
Db	584	AGGATATCTGGTCTTGTGTGTGTGTCGCCGAGGGGCCAATGATGAACTCAATATCTACT	643
QY	569	TCAATTCGAGCGGTGCCAGCTACAGCCGAGAGAGATGTCGGGGGTCCTCTCTCGTGGC	628
Db	644	TCAACTGACTGACTTGAACCCAGCCGTGAGCGCTGTGGGTGCCCTTTTCTCGCTGTG	703
QY	629	TGCCAGATCCTGGCCAAAAGTTGTGAACACACAGTGTGATATGATGATGACATTCAGC	688
Db	704	TTAGGAGACCCAGCGAAGATGTCTCAATATCCAGTGTGGCTATGACATCGACTCAAC	763
QY	689	TGAAGACGAATGGATGATGATGATCTTACAGAAAAGCTGTGATCCAGCGGTGGAAAGCT	748
Db	764	TGGACTCTGACACGAGGGGCTTCATCTTACACAAAGGCTGTGTGGCGCATTTGAGAACT	823
QY	749	GGCCTCCCGGACATTTTACATGTGTGGCTGGCGTTCATCGCCATCTGCTGTTGCAGA	808
Db	824	GGCTACAGACACATCTGATCGTGGGCGCGGGGCTTGTGTGGGATGAGCCCTCTCCAGA	883
QY	809	TATTTGGCATCTTCTCTGGCAAGACGCGTATCTAGACATGAGGACGATGAAGGCC	864
Db	884	TATTTGGCTCTGCTGGCCCAAGACCTGTGATGATCATGAAGCGGTGAAGGCC	939
RESULT_13			
LOCUS	BC010405	2426 bp	mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, similar to RIKEN cDNA 2210021G21 gene, clone		
ACCESSION	MGC:14859 IMAGE:3621871, mRNA, complete cds.		
VERSION	BC010405		
KEYWORDS	BC010405.1 GI:14714540		
SOURCE	MGC.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 2426)		
JOURNAL	Strausberg, R.		
	Direct Submission		
	Submitted (09-JUL-2001) National Institutes of Health, Mammalian		
	Gene Collection (MGC) Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk		
	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Institute for Systems Biology		
	<a href="http://www.systemsbio.org">http://www.systemsbio.org</a>		
	contact: <a href="mailto:amadane@systemsbiology.org">amadane@systemsbiology.org</a>		
	Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia		
	Greene, Mark Ketteman and Anuradha Madan		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 24 Row k Column: 15

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 10440362..

FEATURES		Location/Qualifiers
source	1.. 2426	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:14859 IMAGE:3621871" /tissue_type="Pancreas, adenocarcinoma" /clone_lib="NH_MGC_39" /lab_host="DH10B-R" /note="Vector: pOTB7"
CDS	61.. 852	/codon_start=1 /product="Similar to RIKEN cDNA 2210021G21 gene" /protein_id="AAH10405.1" /db_xref="GI:14714541" /translation="MPGKHQHFDEPEVCCCKYELGFNIVFWLGALEFLAIGLAWMC EGLVNSISALDLGIDPPEVVGWGSVYGFAGCAIRENTFLKPSYVLGAL IFPLILATGILAEVPMKIMRIDOLIFINNVAWRDIDIMOLIDPAQWYSCGAR PDMNUNITRYFNCTDLNPSREKRCYPFCCYCRDPAEDVIMOGGYDRKLEEGGF HTRKEGVGFEMKLDNLIVAGVPMGALILQIFGICLAONLEQME"
BASE COUNT	507 a 662 c 716 g 541 t	
ORIGIN		
Query Match	24.9% Score 345.2; DB 9; Length 2426;	
Best Local Similarity	66.6% Pred. No. 1,2e-68;	
Matches 494; Conservative	0; Mismatches 248; Indels 0; Gaps 0	
95	AAGTCAGCTGCTGGTACAAAGTACCTCTTTTACGTACAAACATCTCTTGGTGGCTG 154	
95	AGGTGCGCTGCTGGGGAATACTCTCTTTGGCTTCAACATCTTCTTGGGGTGTGG 154	
155	GAGTTGCTTCTTGGAGTGGGGGCTGGGCATGGAGCAAAAGGATGCTGTCCGACC 214	
155	GAGCCCTGTCTGGCTATAGGCGCTTGGGCGCTGGGGGTGAGAAAGGCGCTTCTCCGAAACA 214	
215	TCACCAAAAGTAGACCCGGATGATGAATGACACCTGTGTGGTGTCTCTGATGATGGTGGCC 274	
215	TCTCAGGCGCTGACAGATCTGGAGGCGCTTGACCCGCTGTGGCTTTGTGGTAGTTGGAG 274	
275	TGTGTATGTTACACCTCGGGTTGCGCGCTGCGTGGGGGCTGTGGGGAGAAATTCGCT 334	
275	GCGTATGCTCGCTCTGGGCTTTGGCTGTGGCTGCAATTTGGGGCGCTCCGGAGAACCTTCC 334	
335	TGCTCAACTTTTGTGGCAGCACTGATCTATCTTCTCTGGAGCTGGCTGTGGCCG 394	
335	TGCTCAAGTTTTCCTGCTGTCTCTGGGTCTATCTTCTCTGGAGCTGGCAACAGGA 394	
395	TGCTGGCCTTCTCTCCAGAGCTGGGTGAGGAGCCGGTCTCCGGAGTTCTTCCAGAGCA 454	
395	TGCTGCTTTTGTCTTCAAGGACGTGGATTCGACAGCTCAACCTTTTCATCAACAACA 454	
455	ACATCAAGTCTACCGGAGAGATATGATGTGAAACCTCATTCACCTCCCTCGAAG 514	
455	ACGTAAAGGCTACCGGAGAGACATTGACCTCAAGAACTCATTTGACTTGTGTAGGAT 514	
515	CTAACAGTGTGTGGCGCATATGGCCCTGAGAGCTGGAGCCTCAACGTTACTTCAATT 574	
515	ACTGTCTTGTGGCGGAGACCCGAGGCCCATGATGAGAACTCAATAATCTACTTCAACT 574	
575	GCAGCGGTGCCAGCTACAGCCGAGAGAAAGTGGGGGTGCCCTTCTCTGCTGCGCGAG 634	
575	GCACTGACTTGAACCCAGCGGGAGGCGTGGCGGGGTGCCCTTCTCTGCTGCGCGAGG 634	
635	ATTCCTGGCAAAAGTTGTGAACACAGTGTGATATGATGTCAAGATTACGCTGAAGA 694	
635	ACCTCGCGGAGAGTCTCTCAACACCAAGTGTGGCTACGACGTCGGGCTCAAACTGGAGC 694	
695	GCAAGTGGATGATGTCATCTTTCACGAAAGGCTGATCCAGGCGCTGGAAAGCTGGCTCC 754	
695	TGGAGAGAGGCTTATCCACACCAAGGCTGCGTGGGCGCAGTTTGAGAGTGGCTGC 754	
755	CGCGAACAATTACATTGTGGCTGGGCTTCAATCGCATCTGGCTGTGGAGATATTGG 814	

```
Db 755 AGNACAACCTGATGTGGTGGCGGAGTCTTTCATGGCATCGCCCTCCTCCAGATCTTGG 814
QY 815 GCATCTCTCTGGCAAGAGCCT 836
Db 815 GCATCTCTCTGGCGGAGACCT 836

RESULT 14
AF121344 3175 bp mRNA linear ROD 01-FEB-2000
LOCUS Mus musculus tetraspanin Tspan-5 (Tspan5) mRNA, complete cds.
DEFINITION AF121344
ACCESSION AF121344
VERSION AF121344.1 GI:6841032
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3175)
AUTHORS Garcia-Frigola,C., de Lecea,L. and Soriano,E.
TITLE Mouse Tspan-5 cDNA cloning
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3175)
AUTHORS Garcia-Frigola,C., de Lecea,L. and Soriano,E.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1999) Dept. of Animal and Plant Cell Biology,
University of Barcelona, Av. Diagonal 645, Barcelona 08028, Spain
FEATURES
SOURCE
1. 3175
/organism="Mus musculus"
/strain="Swiss Webster/NIH"
/db_xref="taxon:10090"
1..3175
/gene="Tspan5"
/gene="Tspan5"
388..1194
/note="member of transmembrane 4 superfamily"
/codon_start=1
/product="tetraspanin Tspan-5"
/protein_id="AA028869.1"
/db_xref="GI:6841033"
/translaction="MSGKHYKGPVSCIKYFIEGFNVIFWFLGITFLGILMANEX
GVLNLSITDIDGDPVWFLVGVGMFLGFPACIALRENTFLKFFVFLGILF
FLFLTAGVLAFFKDKIDQLYEFINNNRAVDIDIDNLIIDFQEWKCCGAGFAD
DMLNINFTCSNASRERKGVPSCTGDPEDVINFQCGDAROKKEVQOQVITYT
KCGVPOFEKWLQDNLITIVGIFIGIALQIFICLAQNLVSDIEAVRSW"
BASE COUNT 667 a 811 c 840 g 857 t
ORIGIN
Query Match 24.5%; Score 340.2; DB 10; Length 3175;
Best Local Similarity 63.8%; Pred. No. 1.6e-67;
Matches 516; Conservative 0; Mismatches 293; Indels 0; Gaps 0;

QY 59 CTCGAGAGATGACATATTATAGATCTACAGCCCAAGGACGCTGCTGATCAATACC 118
Db 380 CCCAGAGATGTCCTGGAGACACTACAGGGCTCTGAGTCAGTCTTTCATCAATACT 439
QY 119 TCCCTTTCAGCTACACATCTCTGCTGAGTGTGCTGAGTGTCTTCTTGGAGTGGC 178
Db 440 TCATTTTGGCTCAATGATATTGTTGTTTGGAAATACGTTCTTGGAAATCGGAC 499
QY 179 TGTGGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 238
Db 500 TGTGGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 559
QY 239 GAATGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 298
Db 560 GCTTGAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
QY 299 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 358
Db 620 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679

QY 359 TCGTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 418
Db 680 TGGGATTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 739
QY 419 GGGTGGAGGACCGGCTTCCGGAGTCTTCTGAGAGCAATCAAGTCTTACCGGAGATA 478
Db 740 GGATCAAGAGACGCTGATTTCTTTATTAACACACATCAGAGCTTACACAGAGATGACA 799
QY 479 TCGATCTCAAAACCTCATGAGCTCCCTTCCAGAAAGCTTAACAGTCTGTGGCATATG 538
Db 800 TTGATCTACAGAACCTCATGAGCTTCTCACCAGGAATATTTGGACATGCTGTGGCTTTTG 859
QY 539 GCCCTGAAGACTGGAGCTTCAACGCTACTTCAATTCAGCGGCTCCAGCTTACAGCCGAG 598
Db 860 GAGCTGATGATGGAACCTTAATATTTTACTTCAATTCAGATTCACATGCAAGCCGAG 919
QY 599 AGAAGTGGGGGGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 658
Db 920 AGCATCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979
QY 659 CACAGTGTGATATGATGTCAGATTCAGTGAAGAGCAAGTGGATGATCTCATCTTCA 718
Db 980 CTCAGTGTGCTATGATGTCAGGAGCAAGAACCAAGATGACCAAGATGATCTTCA 1039
QY 719 CGAAGGCTCATTCAGAGGCTGGAAGCTGCTCCCGGAGACATTTACATTTGGCTG 778
Db 1040 CAAAAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1099
QY 779 GCGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838
Db 1100 GTATTTTCATAGGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
QY 839 TCTGACATCGAGGAGTGAAGCCGCGC 867
Db 1160 TGAAGTGAATGAAGCTGTCAGGCTAGC 1188

RESULT 15
AF065389 1405 bp mRNA linear PRI 28-APR-2000
LOCUS Homo sapiens tetraspan NET-4 mRNA, complete cds.
DEFINITION AF065389
ACCESSION AF065389
VERSION AF065389.1 GI:3152702
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1405)
AUTHORS Serru,V., Dessen,P., Bouchelx,C. and Rubinstein,E.
TITLE Sequence and expression of seven new tetraspans
JOURNAL Blochm. Biophys. Acta 1478 (1), 159-163 (2000)
MEDLINE 20185353
PUBMED 10719184
REFERENCE 2 (bases 1 to 1405)
AUTHORS Rubinstein,E., Serru,V. and Bouchelx,C.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-1998) INSERM U268, 14 av Paul Vaillant Couturier,
Villejuif 94807, France
FEATURES
SOURCE
1..1405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE ID 219547"
348..1154
/note="similar to Tspan-5; TM4SP"
/codon_start=1
/product="tetraspan NET-4"
/protein_id="AAC17120.1"
/db_xref="GI:3152703"
/translaction="MSGKHYKGPVSCIKYFIEGFNVIFWFLGITFLGILMANEX
GVLNLSITDIDGDPVWFLVGVGMFLGFPACIALRENTFLKFFSVFLGILF
```

FILETAGVLAIEVKDWMIKDQLYEFINNIRAYRDDIDLONLIDFTQEFWCCGAFAD  
DMNLNIFNCTDSDNSARECSRESCREINTCGSDIAROKPEVDQIVITIT  
KCCVPOFEKWLQDNLIIYAGIFIGIALIIFIGICLAONLIVSDIEAVRASW"  
BASE COUNT 309 a 357 c 387 g 352 t  
ORIGIN

Query Match 24.3%; Score 337.4; DB 9; Length 1405;

Best Local Similarity 62.8%; Pred. No. 7.1e-67;

Matches 524; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

QY 33 GCGGCGCGCGCGGATCTCTCTCGAGAGATGCACTATTATAGATACTCTAAACC 92  
DB 314 GCTGCCGGGCAAGTCTTTCATGAACCCAGAGAGATGTCGGGAACACTACAGGGTCC 373  
QY 93 CAAAGTCAGCTGCTGTAGTACAGTACTCTTTTCAGCTACACATCATCTTGTGGC 152  
DB 374 TGAAGTCAGTTGTGATCAATATCTTCAATTTGGCTTCAATGTCATATTTGGTTTTT 433  
QY 153 TGGAGTGTCTCTCTTGTGATCGGGCTGTGGGATGAGCGAAGAGGTGCTGCCGA 212  
DB 434 GCGAATTAACATTTCTTGAATTTGCACTGTGGCATGGAATGAAGAGAGTCTGTCAA 493  
QY 213 CCTCACCAGATGACCCGAGTGCATGCAATGACCCCTGTGTGCTGTGCTGTGAG 272  
DB 494 CATCTTCCATCACCATCTCGCGGCTTTGAOCAGTTTGGCTCTTCTGTGGTGG 553  
QY 273 CGTGTGATGTTCCCTCGGGGTTCCGGGCTGTGGGCTGTGGGAGATATCTG 332  
DB 554 AGGAGTGTATGTTCAATTTGGGATTTGCAAGGATGATGAGCGCTACGGGAAAAACATTT 613  
QY 333 CTTCGTAACATTTCTTCTGTGGCACCATGTCATCTTCTTCCTGAGGCTGTGGC 392  
DB 614 CTTCTCAAGTTTCTTCTGTGTCTGTGGATTTATTTCTTCTGTGAGCTCACTGCCG 673  
QY 393 CGTCTGCGCTTCTCTGTCCAGAGTGGGTGAGGACCGGTTCCGGAGTTCTTCAGAG 452  
DB 674 AGTTCTAGCATTTGTTTCAAGAGTGTATCAAGACCATGATTTCTTTAATAACAA 733  
QY 453 CAACATCAAGTCTTACCGGAGCATGATGATCTGCAAAACCTCATGACTCCCTTCGAA 512  
DB 734 CAACATCAAGCATATCGGATGATGATTTGCAAAACCTCATGACTCCACCCAGGA 793  
QY 513 AGCTAACAGATGCTGTGGCATGATGAGGCTGAGACTGAGACTCAACGTCCTACTCAA 572  
DB 794 ATATTGCAAGTGTGTGGGCTTTTGGAGCTGATGATTTGAACTAATAATTTACTTCA 853  
QY 573 TTGACGCGGTGCGAGCTACAGCCGAGAGAGTGGGGTCCCTTCTCTGTGCTGCC 632  
DB 854 TTGCACAGATTTCAATGCAAGTGCAGAGCATGTGGGCTTCCATTTCTCTGCTGCACTAA 913  
QY 633 AGATTCCTGCCAAGATTGTGAACACAGTGTGATGATGATGATGATGATGATGATGAA 692  
DB 914 AGATCCCGAGAGATGTCATACACTGAGTGTGCTATGATGATGATGATGATGATGATGAA 973  
QY 693 GAGCAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAA 752  
DB 974 AGTTGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAA 1033  
QY 753 CCGCGGGAACATTTTACATGCTGTGCTGTCTTCAATGCCCATCTGCTGTGCAATAT 812  
DB 1034 GCAGGACAAATTTAAACCATGCTGTGCTGTATTTTCAATGAGCATTTGATGCTGCAATAT 1093  
QY 813 TGGCATCTTCTGCGAG 867  
DB 1094 TGGGATATGCTGTGGCCAGAAATTTGTTAGCATATGCAAGTGTCAAGGCGGAGC 1148

Search completed: December 19, 2002, 04:48:40  
Job time : 2928 secs



```

; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 243
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Mouse
US-09-188-930-243

Query Match
Best Local Similarity 5.7%; Score 79.8; DB 3; Length 399;
Matches 117; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 95 AAGTACGCTGCTGACAGTACCTCTTTCAGCTACACATCATCTTGGTGGCTG 154
    ||||| || || || || || || || || || || || || || || || ||
DB 11 AAGTCAGTTGTGCATCAAAATCTTCATTTTGGCTTCATATTTTGGTTTGG 70
    || || || || || || || || || || || || || || || || || || ||
QY 155 GAGTGTCTCTCTGAGTGGGCTGTGGGATGAGGAAAGGCTGTGCTCCGACC 214
    || || || || || || || || || || || || || || || || || || ||
DB 71 GAATAACGTTCTGTGAATCGGACTGTGGCGTGAATGAAAAGTGTCTCTCCACA 130
    || || || || || || || || || || || || || || || || || || ||
QY 215 TCACCAAGTGAACCCGATGATGGAATGACCTGTGTGCTGCTCGATGATGAGGC 273
    || || || || || || || || || || || || || || || || || || ||
DB 131 TCTGTCCATCAACGACCTCGGTGGCTTGAACCCAGTGTGGCTTCTCTGAGTGGCC 189
    || || || || || || || || || || || || || || || || || || ||

RESULT 3
US-08-254-493-2
; Sequence 2, Application US/08254493
; Patent No. 5439886
; GENERAL INFORMATION:
; APPLICANT: IKEYAMA, SHUICHI
; APPLICANT: KOYAMA, MASARU
; APPLICANT: MIYAKE, MASAYUKI
; APPLICANT: SENOO, MASAHARU
; TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
; TITLE OF INVENTION: PRODUCTION THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/254,493
; FILING DATE:
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/865552
; FILING DATE: 09-APR-1992
; APPLICATION NUMBER: JP 079996-1991
; FILING DATE: 12-APR-1991
; APPLICATION NUMBER: JP 085396-1991
; FILING DATE: 17-APR-1991
; APPLICATION NUMBER: JP 022321-1992
; FILING DATE: 07-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
```

```

; REFERENCE/DOCKET NUMBER: 41777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELETYPE: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 687 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HUMAN
; CELL TYPE: BREAST CARCINOMA
; CELL LINE: ZR-75-1
US-08-254-493-2

Query Match
Best Local Similarity 5.0%; Score 69.8; DB 1; Length 687;
Matches 199; Conservative 0; Mismatches 167; Indels 9; Gaps 2;

QY 96 AGTCAGCTGCTGTGACAGTACCTCTTTCAGCTACACATCATCTTGTGCTGG 155
    || || || || || || || || || || || || || || || || || || ||
DB 15 AGGACCAAGATGATCAATATACCTGCTGTGGATTTAATTCATCTGTGCTCCGG 74
    || || || || || || || || || || || || || || || || || || ||
QY 156 AGTTGTCTCTTGTGAGTGGGCTGTGGGATGAGCGAAGAGGTGCTGTCCAGCT 215
    || || || || || || || || || || || || || || || || || || ||
DB 75 GATTGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 134
    || || || || || || || || || || || || || || || || || || ||
QY 216 CACCAAGTGAACCCGATGATG-----GAATGACCTGTGTGCTGTCTGAT--- 266
    || || || || || || || || || || || || || || || || || || ||
DB 135 CGACCAAGAACTAATATATATATATATATATATATATATATATATATATATAT 194
    || || || || || || || || || || || || || || || || || || ||
QY 267 GGTGGCGGTGTGATGTTACCTGTGGGCTGCCGCTGCTGCTGCTGCTGCTGCTG 326
    || || || || || || || || || || || || || || || || || || ||
DB 195 AGCCGGCGCCCTCATGATGCTGTGGGCTCTGCGGCTGCTGCGGAGATC 254
    || || || || || || || || || || || || || || || || || || ||
QY 327 TATCTGCTCTCACTTTTCTGTGTGACCATGCTGCTCATCTCTCTGTGAGCTGG 386
    || || || || || || || || || || || || || || || || || || ||
DB 255 CCACTGATGCTGTGGAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 314
    || || || || || || || || || || || || || || || || || || ||
QY 387 TGTGGCGGTGCTGCTCTCTGTTCCAGACTGGTGTGAGGAGCGGTCCGGAGTTCT 446
    || || || || || || || || || || || || || || || || || || ||
DB 315 TCGCGCATCTGTGGGATATCCCAAGATGAGTGATTAAGGAAGTCCAGAGTTTAA 374
    || || || || || || || || || || || || || || || || || || ||
QY 447 CGAGAGCAACATCAA 461
    || || || || || || || || || || || || || || || || || || ||
DB 375 CAAGACACCTACAA 389
    || || || || || || || || || || || || || || || || || || ||

RESULT 4
US-08-408-222B-2
; Sequence 2, Application US/08408222B
; Patent No. 5776727
; GENERAL INFORMATION:
; APPLICANT: Ikeyama, Shuichi
; APPLICANT: Koyama, Masaru
; APPLICANT: Miyake, Masayuki
; APPLICANT: Senoo, Masaharu
; TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,222B
FILING DATE: 22-MAR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,493
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-079996-1991
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-085396-1991
FILING DATE: 14-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-022321-1992
FILING DATE: 07-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41777-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: breast carcinoma
CELL LINE: ZR-75-1
US-08-408-222B-2

Query Match
Best Local Similarity 5.0%; Score 69.8; DB 1; Length 687;
Matches 199; Conservative 0; Mismatches 167; Indels 9; Gaps 2;

DB 96 AGTCAGCTGCTGTACAGTACCTCTTTTCAGCTACACATCATCTCTGTTGGCTGG 155
15 AGGACCAAGTGCATTAATACCTGCTGTCGATTTAACTTCACTTCTGCGCTTGC 74
OY 156 AGTTGCTTCTTGAGAGTGGGCTGGGCGATGAGCGAAAGGGTGTGCTGCCACT 215
75 GATTGCTGCTCCATGCTGAGTATGGCTCCGATTCGACTCGACAGCAAGAGCATCTT 134
OY 216 CACCAAGTGCAGCGATGCATG-----GATCGACCTGTGCTGCTGCTGAT--- 266
135 CGAGCAAGAACTAATATATATATTCAGCTCTACACAGAGTCTATATTTGTATCGG 194
OY 267 GGTGGCGGTGTATGTTTCACTCGGGGTTGCGGCGTGGGCTGCGGGAGAA 326
195 AGCGGCGGCTCATGATGCTGTGGGCTTCGCGGCTGCTGCGGCTGTCGAGGAGTC 254
OY 327 TATCTGCTGTCAACTTTTCTGTGTCGACCATGTCCTATCTTCTCTGAGCTGGC 386
255 CCACTGATGCTGGAGCTGTTCTGCTTCCTGCTTCCTGATATGTCCTTAATAAGC 314
DB 387 TGTGGCGTGTGCTCTCTGTTCCAGGACTGGGTGAGGAGCGGTTCCGGAGTCTT 446
315 TGGCGCATCTGGGATATATCCACAGATGAGTGATTAAGAGTCCAGAGTTTNA 374
OY 447 CGAGAGCAACATCAA 461
DB 375 CAAGACACCTACAA 389
```

```
RESULT 5
US-08-408-222B-3
Sequence 3, Application US/08408222B
Patent No. 5776727
GENERAL INFORMATION:
APPLICANT: Ikeyama, Shuichi
APPLICANT: Koyama, Masaru
APPLICANT: Miyake, Masayuki
APPLICANT: Senoo, Masaharu
TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronstein, Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,222B
FILING DATE: 22-MAR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,493
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-079996-1991
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-085396-1991
FILING DATE: 14-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-022321-1992
FILING DATE: 07-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41777-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: breast carcinoma
CELL LINE: ZR-75-1
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 112...795
OTHER INFORMATION: E Mat peptide
US-08-408-222B-3

Query Match
Best Local Similarity 5.0%; Score 69.8; DB 1; Length 1120;
Matches 199; Conservative 0; Mismatches 167; Indels 9; Gaps 2;

OY 96 AGTCAGCTGCTGTACAGTACCTCTTTTCAGCTACACATCATCTTCTGTTGGCTGG 155
```

Db 126 AGCACCAGTGCATCAAAATACCTGCTGTTCGATTTAACTTCATCTTCGGCTGCCG 185  
156 AGTTGCTTCTTGAGTGGGGCTGTGGGANTGAGCGAAAGGGTGTCTCCGACCT 215  
Db 186 GATTGCTGTCTTCCATTGGACTATGGCTCCGATTCGACTTCAGACCAAGAGCATTT 245  
Qy 216 CACCAAGTGTACCCGATGATG-----GAATGACCTGTGTGTCTGTCTGTAT--- 266  
Db 246 CGAGCAAGAACTAATAAATAATATTCAGCTTCTACAGAGAGTCTATATCTGTATGCG 305  
Qy 267 GGTGGCGTGTGTATTTACCTTGCGGGTTCGCGGGCTGTGGGGCTCTCGCGAGAA 326  
Db 306 AGCGGGCGGCTTATGCTGTGGGCTTCCTGGGGCTCTCGGGGCTGTGACAGAGTC 365  
Qy 327 TATCTGCTGCTCACTTTTCTGTGGCAGCATGCTGCTCATCTTCTTCCTGAGCTGCG 386  
Db 366 CAGTGCATGCTGGGAGCTGTTCTTGCGCTTCTTGATGATTCGCGCATTTGAATAGC 425  
Qy 387 TGTGGCGTGTGGGCTTCTTCTTCAGAGACTGTGAGGAGCGGTTCCGGAGTTCTT 446  
Db 426 TGGGGCATCTGGGATATTCACAGAGATGAGTGTATGAAGATCCAGAGATTTTA 485  
Qy 447 CGAGAGCAATCA 461  
Db 486 CAAGGACACTACAA 500

## RESULT 6

US-09-333-599-5

Sequence 5, Application US/09333599

Patent No. 6245898

GENERAL INFORMATION:

APPLICANT: Testa, Jacqueline E.

APPLICANT: Quigley, James P.

APPLICANT: Seandel, Marco

TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS

FILE REFERENCE: SUNY

CURRENT APPLICATION NUMBER: US/09/333,599

CURRENT FILING DATE: 1999-06-15

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5

LENGTH: 827

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (57)..(815)

US-09-333-599-5

## Query Match

4.9%; Score 68.6; DB 4; Length 827;

Best Local Similarity 48.4%; Pred. No. 4e-09;

Matches 222; Conservative 0; Mismatches 224; Indels 3; Gaps 1;

Qy 97 GTCAGCTCTGTGATACAGTACCTCTTTTACGCTACACATCATCTTGTGGCTGA 156  
Db 90 GGCACCTGTTGCTCAAGTACCTGCTGTTTACCTACATCTGCTGCTGGCTGGCGC 149  
Qy 157 GTTGTCTTCTTGAGTGGGGCTGTGGGATGAGCGAAAGGTTGTCTCCGACCTC 216  
Db 150 CTGGCTGTATGGAGTGGGATGCTGAGCGCTGCCCTCAAGAGTGTACTACATGAGCTG 209  
Qy 217 A---CCAAGTACCCGAGTGCATGATGAGATCGACCTGTGGTGGTGGTGGTGGG 273  
Db 210 CTGGCTGTAGGACACTTACTGTGGCAGACCTACATCTCTGTGGTGGGCTGTGTC 269  
Qy 274 GTGTGATGTTCACCTGTGGGTTGCGGCTGCGTGGGGCTCTGCGGAGATATCTGC 333  
Db 270 ATGTGATGTGGGCTGTGGGCTGTGCGGCTTCAAGGAGCGTGGAACTGTGCGC 329  
Qy 334 TTGCTCAACTTTTCTGTGGCAGCATGCTGCTCATCTTCTTCTGAGCTGGCTG 393

Db 330 CTGTACTTCACTGCTGCTCTCTATCATCTTTTGTGTGGAGATCATGCTGTATCTGCC 389  
Qy 394 GTGTGGCTTCTCTGTTCCAGAGTGGGTGAGAGGAGCGGTTCCGGAGTTCTTGTGAGAGC 453  
Db 390 TACGCTTACTACAGACACTGATACAGAGGATGTAAGAGAACTGTAAGAGACCATGACC 449  
Qy 454 AACATCAAGTCTTACCGGAGAGATGTGATGTGCAAAAACCTCATGACTCTCTTACAAA 513  
Db 450 AAGCGTACACAGCGCGGCTATGAGGCTGTGACACAGCGCTGTGACACAGCTGCACAG 509  
Qy 514 GCTAACAGTCTGTGGGCTATGAGCCCTGGAAGACTGG 552  
Db 510 GAGTTCCACTGCTGTGGCAGCAACACTCACAGAGACTGG 548

## RESULT 7

US-09-333-599-1

Sequence 1, Application US/09333599

Patent No. 6245898

GENERAL INFORMATION:

APPLICANT: Testa, Jacqueline E.

APPLICANT: Quigley, James P.

APPLICANT: Seandel, Marco

TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS

FILE REFERENCE: SUNY

CURRENT APPLICATION NUMBER: US/09/333,599

CURRENT FILING DATE: 1999-06-15

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 870

TYPE: DNA

ORGANISM: Homo sapiens

US-09-333-599-1

## Query Match

4.9%; Score 68.6; DB 4; Length 870;

Best Local Similarity 48.4%; Pred. No. 4.1e-09;

Matches 222; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

Qy 97 GTCAGCTCTGTGATACAGTACCTCTTTTACGCTACACATCATCTTGTGGCTGA 156  
Db 134 GGCACCTGTTGCTCAAGTACCTGCTGTTTACCTACATCTGCTGCTGGCTGGCG 193  
Qy 157 GTTGTCTTCTTGAGTGGGGCTGTGGGATGAGCGAAAGGTTGTCTCCGACCTC 216  
Db 194 CTGGCTGTATGGAGTGGGATGCTGAGCGTGGCCCTCAAGAGTACTACATCACCTG 253  
Qy 217 A---CCAAGTACCCGAGTGCATGATGAGATCGACCTGTGGTGGTGGTGGTGGG 273  
Db 254 CTGGCTGTAGGACACTTACTGTGGCAGACCTTACATCTCTGTGGTGGGCTGTGTC 313  
Qy 274 GTGTGATGTTCACCTGTGGGTTGCGGCTGCGGCTGTGGGCTGTGCGGAGATATCTGC 333  
Db 314 ATGTGATGTGGGCTGTGGGCTGTGCGGCTTCAAGAGAGTGGAACTGTGCGC 373  
Qy 334 TTGCTCAACTTTTCTGTGGCAGCATGCTGCTCATCTTCTTCTGAGCTGGCTG 393  
Db 374 CTGTACTTCACTGCTGCTGCTCATCATCTTTTGTGGAGATCATGCTGTGATCTCGGC 433  
Qy 394 GTGCTGCTCTCTGTTCCAGAGCTGGGTGAGGAGACCGGTTCCGGAGTTTTCGAGAGC 453  
Db 434 TACGCTTACTACAGACAGTGAACAGGAGCTCAAGAGAACTGTAAGAGACCATGACC 493  
Qy 454 AACATCAAGTCTTACCGGAGAGATGTGATGTGCAAAAACCTCATGACTCTCTTACAAA 513  
Db 494 AAGCGTACACAGCGCGGCTATGAGGCTGTGACACAGCGCTGTGACACAGCTGCACAG 553  
Qy 514 GCTAACAGTCTGTGGGCTATGAGCCCTGGAAGACTGG 552  
Db 554 GAGTTCCACTGCTGTGGCAGCAACACTCACAGAGACTGG 592



QY 96 AGTACGCGCGGGAGAACAGATCCCTTTTCAGCTACACAATCATCTTCTGTGGCTGG 155  
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 126 AGGCACCAAGTGACATGCCAAATACCCTGCTTCCGGAATTAACTTCACTTCTGGCTGGCG 185  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 156 AGTTGCTCTTCCTTAGAGTCCGGGCTGTGGGCGATGGAGCGAAAAGGNGTGTGCCACT 215  
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 186 GATTGCTGTCTCCTTCATTGTGGACTATGAGTCCGATTCGACTCTCAGACCAAGAGCATCTT 245  
   | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |  
 QY 216 CACCAAGTACCCGGATGCATG-----GANTGACCCCTGTGTGCTGGTCCCAT --- 266  
   | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |  
 Db 246 CGAGCAGAAGAACTAATAATATAATTAATTCACAGCTTCTTACACAGAGATCTAATAATTCAGACGG 305  
   | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |  
 QY 267 GGTGGAGCGGTGATGTATTTCAACCCCTGGGGGTTCCGCGCTGCGTGGGCTCTCGGGAGAA 326  
   | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |  
 Db 306 AGCGCGGCGCCCTCATGATGCTGAGTGGGCTTCTGAGGCTGCTGCGGGGCTCTGGCAGGATC 365  
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 327 TATCTGCTTCTCAACTTTTTTCTGTGGCACCATCTGCTCATCTTCTCTCGAGCTGGC 386  
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 366 CCAGTGCATCTGGGAGACTGTCTTCGCGCTTCCTTGCTGATATATTCCTCCCAATGAAATAGC 425  
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 387 TGTCGCGCTGCTGGACCTTCCTGTTCCAAGACTGGGTGAGGAGCCGTTCCGGGAGTCTT 448  
   ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 426 TCGCGCCATCTGGGGATATTCACACAAGATGAGTGTATTAAGSAAATCCACGAGATTTTA 485  
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 QY 447 CGAGAGCAACATCAA 461  
   ||| ||| ||| |||  
 Db 486 CAAGAGCACCTACAA 500  
   ||| ||| ||| |||  
 RESULT 9  
 US-08-705-771-8  
   Sequence 8, Application US/08705771  
   Patent No. 6054289  
 GENERAL INFORMATION:  
   APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
   APPLICANT: Jian Ni and Jing-Shan Hu  
   TITLE OF INVENTION: Human Genes, Sequences and  
   TITLE OF INVENTION: Expression Products  
   NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESSES:  
   ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN,  
   ADDRESSSEE: CECCHI, STEWART & OLSTEIN  
   STREET: 6 BECKER FARM ROAD  
   CITY: ROSELAND  
   STATE: NEW JERSEY  
   COUNTRY: USA  
   ZIP: 07068  
 COMPUTER READABLE FORM:  
   MEDIUM TYPE: 3.5 INCH DISKETTE  
   COMPUTER: IBM PS/2  
   OPERATING SYSTEM: MS-DOS  
   SOFTWARE: WORD PERFECT 5.1  
 CURRENT APPLICATION DATA:  
   APPLICATION NUMBER: US/08/705,771  
   FILING DATE: August 30, 1996  
   CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
   NAME: MULLINS, J.G.  
   REGISTRATION NUMBER: 33,073  
   REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)  
 TELECOMMUNICATION INFORMATION:  
   TELEPHONE: 973-994-1700  
   TELEFAX: 973-994-1744  
 INFORMATION FOR SEQ ID NO: 8:  
   SEQUENCE CHARACTERISTICS:  
     LENGTH: 1344 base pairs  
     TYPE: nucleic acid  
     STRANDEDNESS: single  
     TOPOLOGY: linear  
     MOLECULE TYPE: DNA  
 US-08-705-771-8

```

; IMMEDIATE SOURCE:
; LIBRARY: MYOMNOT01
; CLONE: 779308
US-08-855-140-2

```

Query Match	4.58;	Score 61.8;	DB 2;	Length 977;
Best Local Similarity	48.08;	Pred. No. 2.9e-07;		
Matches 248;	Conservative	0;	Mismatches 257;	Indels 12; Gaps 2

102 CTGCTGCTACAGTACCTCTTTTCAGCTACACAATCATCTTCTGCTGGCTGGAGTTGT 16

162 CTTCCCTGAGATCGGGCTGTGGCCATGGAGCGAAAGGGTGTGCTGTCCGACCTACCA 22

b 199 GCTGCTGGGAGTGGGCATCTGACTCTCCGTGTCCCAAGGCACACTTTCACCTTCTCC -- 25

257 ----CCAGCTTCCCTTCGTTCTCTGCAGCCACCTGGTCATCGCCATAGGCACCATTTCT 31

282 GTTACCTGGGGTTCGCCGCTGCGGCTCTCGGGAGATATCTGCTTGCCTCA 34

342 CTTTTCCTGTCGACCATCGTCTCATCTTCTTCTCTGAGCTGGCTGTGGCCGCTGCC 40

373 CAGCTTTTCATCGTCCTGTTGGTCACCCCTCCTAGCAGAGCGTGAATCTTACCCATCCCTT 43

433 CTTTGTCATATGACCAAGGTGAACGAGAACGCCAAGAAGACTGAAGGAAGGCTGCT 49

493 GCTGTCCACACCGAGAACCTGGGGCTGAGAACCTCTGAACATCATCCAGCTGA 55

516 TAACCACTGCTGTGGCCATATGGCCCTGAAGACTGGGACCTCAACGCTACTTCAATTG 57

576 CAGGCGGGCCACGCTTACAGCCTCGAGAGAAAGTCGCGGCCTC 612

576 CAGCGTGCCAGCTACAGCCGAGAGAGTCCGGGGTC 612

613 GGTTCGCCGACCGCTGCTGCATGGAGACTCCAGGGC 649

013 GGATCCGACCGGCTGCTGCAAGGAACTCCAGGGG 649

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847

RESULT 11  
5-08-807-044-2

Sequence 2, Application US/08807044

Sequence 2, Application 05/0880/044  
Patent No. 5863735

Patent No. 2003/05  
GENERAL INFORMATION:  
APPLICANT: HILLMAN  
TENNIS FOR I

APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surva K.

APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN TRANSMEMBRANE 4 SUPERFAMILY

TITLE OF INVENTION: NOVEL HUMAN TRANSMEMBRANE 4 SUPERFAMILY  
NUMBER OF SEQUENCES: 4

NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Institute of Mathematics, University of

ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive

STREET: 3174 Porter Drive  
CITY: Palo Alto

CITY: Palo Alto  
STATE: CA

STATE: CA  
COUNTRY: USA  
STD: 01204

COMPUTER READABLE FORM:  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

```
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
```

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/807,044

APPLICATION NUMBER: US/08/807,044  
FILING DATE: Herewith

FILING DATE: Herewith  
CLASSIFICATION: 530

CLASSIFICATION: 530  
PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

APPLICATION NUMBER:

Query Match	4.48;	Score 61.4;	DB 2;	Length 1151;
Best Local Similarity	51.9%;	Pred. No. 4e-07;		
Matches 162; Conservative	1;	Mismatches 143;	Indels 6;	Gaps 1

RESULT 12

```

1 APPLICATION NUMBER: US/08/430,225A
2 FILING DATE: 28-APR-1995
3 CLASSIFICATION: 514
4 ATTORNEY/AGENT INFORMATION:
5 NAME: RICHARD W. BORK
6 REGISTRATION NUMBER: 36,459
7 REFERENCE/DOCKET NUMBER: 2076-4172
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: (212) 758-4800
10 TELEFAX: (212) 751-6849
11 TELEEX: 421792
12 INFORMATION FOR SEQ ID NO: 19:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 1624 base pairs
15 TYPE: nucleic acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18
19 US-08-430-225A-19
20
21 Query Match 3.9%; Score 54.4; DB 4; Length 1624;
22 Best Local Similarity 47.6%; Pred. No. 3.6e-05;
23 Matches 160; Conservative 0; Mismatches 176; Indels 0; Gaps 0.
24
25 QY 111 CAAGTACCTTCCTTTGACGTACACATCATCTTCGTGGCTGGAGTTCTTCCTGG 170
26   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 Db 192 CAAATACCTTCCTTCCTTCCTTCATCATGTATCTTTATCCTGGGGCAGTATGCTGG 251
28
29 QY 171 AGTCGGGCTGTGGGATGAGGCAAAAGGTTGTCTGTCCGACTACCAAATGACCG 230
30   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
31 Db 252 CTTCGGGGTGTGGATCTGGCCGACGAAGAGAGTTTCATCTCTCTGCAAACTCCTC 311
32
33 QY 231 GATGCATGGAATGACACCTGTGTGCTGTCCTGATGGTGGCGTGTGATGTTCAACCT 290
34   | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35 Db 312 CAGCTGGCTTAGAATGGGGGCTATGTCTTATCGGCTGGGGCAGTACTATGCTCAT 371
36
37 QY 291 GGGGTTGCCCGGCTGCGTGGGGCTCTGCGGAGAAATATCTGTTGCTCAACTTTTCTG 350
38   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
39 Db 372 GGGCTTCCTGGGCTGATCGGCGCCGTCAACAGAGTCCGCTGCTGCTGGGCGCTGACTT 431
40
41 QY 351 TGGCACCATCGTGCATCATCTTCTCTGAGACTGGCTGTGGCGGTGTGTGGCCCTTCTGTT 410
42   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43 Db 432 TGCCTTCCTGCTCTGATCTCATTTGCCCCAGGTACGCGGGGCCCTCTTACTTCAA 491
44
45 QY 411 CCAAGACTGGGTGAGGGACCGGTTCCGGAGTCTT 446
46   | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 Db 492 CATGGCAAGCTGAACAGAGAGATGGGGGCGATCGT 527
48
49 RESULT 13
50 US-08-807-044-4
51 ; Sequence 4, Application US/08807044
52 ; Patent No. 5863735
53 ;
54 GENERAL INFORMATION:
55 APPLICANT: Hillman, Jennifer L.
56 APPLICANT: Goli, Surya K.
57 TITLE OF INVENTION: NOVEL HUMAN TRANSMEMBRANE 4 SUPERFAMILY
58 NUMBER OF SEQUENCES: 4
59 CORRESPONDENCE ADDRESS:
60 ADDRESSEE: Incyte Pharmaceuticals, Inc.
61 STREET: 3174 Porter Drive
62 CITY: Palo Alto
63 STATE: CA
64 COUNTRY: USA
65 ZIP: 94304
66
67 COMPUTER READABLE FORM:
68 MEDIUM TYPE: Diskette
69 COMPUTER: IBM Compatible
70 OPERATING SYSTEM: DOS
71 SOFTWARE: FASTSEQ for Windows Version 2.0
72 CURRENT APPLICATION DATA:
73 APPLICATION NUMBER: US/08/807,044
74 FILING DATE: Herewith
75 CLASSIFICATION: 530

```



STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/020,956  
FILING DATE: 09-FEB-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.427C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 801 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-020-956-16



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2002, 04:49:08 ; Search time 169 Seconds

(without alignments)  
3259.398 Million cell updates/sec

Title: US-09-905-674-1

Perfect score: 1388  
Sequence: 1 ctccctcgagccgagccggc.....ggagagcgagacttgcc 1388

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 356696 seqs, 198428768 residues

Total number of hits satisfying chosen parameters: 713392

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_MA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PC1\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PC1\_NEW\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	362	26.1	813	9	US-09-934-268-3
2	362	26.1	3184	9	US-09-934-268-1
3	351.8	25.3	2538	9	US-09-972-970-2
4	337.4	24.3	1655	10	US-09-925-299-205
5	245.8	17.7	446	10	US-09-783-590-10353
6	230	16.6	864	9	US-09-875-440-1
7	229	16.5	1110	10	US-09-729-674-169
8	165.2	11.9	545	10	US-09-764-873-12
9	122.8	8.8	1778	12	US-10-052-586-277
10	113.4	8.2	368	12	US-10-042-417-37
11	108.2	7.8	288	10	US-09-764-873-14
12	72	5.2	441	10	US-09-960-352-1269
13	69.8	5.0	547	10	US-09-895-828-318
14	69.8	5.0	1192	10	US-09-860-107-2351
15	69.8	5.0	1219	10	US-09-822-830A-421
16	68.6	5.0	1350	10	US-09-925-301-539
17	68.6	4.9	1545	10	US-09-925-301-392
18	68.2	4.9	551	10	US-09-895-828-400
19	66.2	4.8	430	9	US-10-040-739-658

20	63.4	4.6	401	10	US-09-960-352-11404	Sequence 11404, A
21	61.8	4.5	1175	10	US-09-823-356-31	Sequence 31, App1
22	58.2	4.2	343	9	US-10-046-935-372	Sequence 372, App
23	58.2	4.2	343	9	US-09-878-178-372	Sequence 372, App
24	58.2	4.2	466	10	US-09-895-828-327	Sequence 327, App
25	58.2	4.2	500	10	US-09-895-828-325	Sequence 325, App
26	58.2	4.2	515	10	US-09-895-828-326	Sequence 326, App
27	58.2	4.2	559	10	US-09-895-828-301	Sequence 301, App
28	58.2	4.2	594	10	US-09-895-828-340	Sequence 340, App
29	58	4.2	473	9	US-09-736-457-1565	Sequence 1565, Ap
30	58	4.2	473	9	US-09-902-941-1565	Sequence 1565, Ap
31	57.2	4.1	481	10	US-09-895-828-328	Sequence 328, App
32	56.6	4.1	320	10	US-09-895-828-338	Sequence 338, App
33	55.6	4.0	413	10	US-09-825-302-313	Sequence 313, App
34	55.6	4.0	572	10	US-09-998-598-563	Sequence 563, App
35	55.6	4.0	690	10	US-09-895-828-430	Sequence 430, App
36	55.6	4.0	873	10	US-09-864-864-279	Sequence 279, App
37	55.2	4.0	308	10	US-09-895-828-336	Sequence 336, App
38	55.2	4.0	555	10	US-09-895-828-337	Sequence 337, App
39	54.8	3.9	416	9	US-09-954-531-440	Sequence 440, App
40	54.8	3.9	450	10	US-09-895-828-335	Sequence 335, App
41	54.4	3.9	1624	10	US-09-795-380-19	Sequence 19, App1
42	54.2	3.9	929	10	US-09-823-356-34	Sequence 34, App1
43	54.2	3.9	1529	10	US-09-799-777-79	Sequence 79, App1
44	54	3.9	368	10	US-09-895-828-337	Sequence 337, App
45	53.2	3.8	353	10	US-09-960-352-6249	Sequence 6249, Ap

#### ALIGNMENTS

RESULT 1  
US-09-934-268-3  
Sequence 3, Application US/09934268  
Patent No. US20020172986A1  
GENERAL INFORMATION:  
APPLICANT: Lelby, Kevin R.  
TITLE OF INVENTION: 23228, A NOVEL HUMAN TETRASPANIN FAMILY  
FILE REFERENCE: 10448-079001  
CURRENT APPLICATION NUMBER: US/09/934,268  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: 60/226,612  
PRIOR FILING DATE: 2000-08-21  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 813  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-934-268-3

Query Match 26.1%; Score 362; DB 9; Length 813;  
Best Local Similarity 66.9%; Pred. No. 3.6e-88;  
Matches 515; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY	95	AAGTCAGCTGCTGTGACAGTACCTCTTTCAGCTCAACATCATCTTGTGCTGCTG	154
DB	35	AGCTGCGCTGCTGCGGGAATCTCTGTTGGCTTCAACATCTTGTGCTGCTG	94
QY	155	GAGTGTCTTCCTTGAGTGGGCTGTGGCANGAGCAAGAGGCTGTCTCCGACC	214
DB	95	GAGCCGCTTCTCGATCGGCTCTGCGGCTGTGGATGGAAGGCTTCTCTCCACA	154
QY	215	TCACCAAGTACCGGATGATGCAACCTGTGTGCTGTCTGTGATGTTGGCG	274
DB	155	TCTCAGGCTGACAGATCTGGAGGCTTGACCCGCTGTGTGTGTGTGTGAG	214
QY	275	TGTGATGTTACACCTGGGCTGCGGCTGCGGCTGCGGAGGAATATGCT	334
DB	215	GCGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	274
QY	335	TGCTCAACTTTTCTGTGTGTCATCTTCTTCTGAGTGTGCTGCTGCTGCTG	394



Db	275	TGCTCAAGTCTTCTGCTCCGTCTCTCGGCTCACTCTTCTCTGAGACTGGCAACAGGCA	334
QY	395	TGCTGAGCCTTCTCTTCCAGAGCTGGGTGAGGACCGGTTCCGGAGATCTTGTGAGAGCA	454
Db	335	TCTCTGGCCTTGTCTTCTCAAGACCTGGATTTGAGAACCAAGCTCAACCTTCTATCAACAGCA	394
QY	455	ACATCAAGTCTCTACCGGAGCATATGATCTGCAAAACCTCATCTACTCCCTTCAGAAAG	514
Db	395	ACGTCAAGGCTTACCGGGAGACATTTGACCTCCAGAACCTCATTTGACTTGGCTGAGGAAAT	454
QY	515	CTAACCAAGTGTGTGGGGCATATGGCCCTTGAGAGCTGGGACCTCAAGCTCATCTCAAT	574
Db	455	ACTGCTCTTGTGTGGAGACCCGAGAGGCCCAATGCTGGAACCTCAATATCTACTTCACT	514
QY	575	GCACCGGTGCCAGCTACAGCCGAGAGAAATGTGGGGGTCCCTTCTCTGCTGGCTGCAG	634
Db	515	GCACGTACTTAACCCCAAGCCGGGAGGGGTGGGGGGTCCCTTCTCTGCTGGCTGCAGG	574
QY	635	ATCTCTGGCAAAAAGTTGTGAAACAACACAGTGTGATATGATGTGAGATTACGTCAAGA	694
Db	575	ACCTCTGGGAGAGTCTCTCAACACCCAGTGTGCTACGAGTCTGGCTCAAACTTGGAGC	634
QY	695	CGAAGTGGAGATGATCCCATCTTTCAGAAAGGCTGCATTCAGAGGCTGGAAAGCTGGCTCC	754
Db	635	TGGAGCAACAGGGCTCTACCTACACCAAGGCTGCTGGGCCAATTGGAGAATGGCTGCC	694
QY	755	CGCGGAACATTTACATTGTGTGGCTGGGCTGCTTCACTACGCGCATCTCGCTGTGGAGATATTG	814
Db	695	AGGACAAACCTGATTTGTGTGGCGGGAGTCTTCAATGGGAGTGGCCCTCTCCATGATCTTTG	754
QY	815	GCATCTTCTGTGGCAAGCAGCTGATCTTCAGACATTCAGAGCAGTGAAGGCC	864
Db	755	GCATCTGCTGTGGCCAGAACCTCGTGTGATGTGCATTCAGAGCAGTGAAGGCC	804

	RESULT 2	
US-09-934-268-1		
: Sequence 1, Application US/09934268		
: Patent No. US20020172986A1		
: GENERAL INFORMATION:		
: APPLICANT: leiby, Kevin R.		
: TITLE OF INVENTION: 23228, A NOVEL HUMAN TETRASPANIN FAMILY		
: TITLE OF INVENTION: MEMBER AND USES THEREOF		
: FILE REFERENCE: 10448-079001		
: CURRENT APPLICATION NUMBER: US/09/934,268		
: CURRENT FILING DATE: 2001-08-21		
: PRIOR APPLICATION NUMBER: 60/226,612		
: PRIOR FILING DATE: 2000-08-21		
: NUMBER OF SEQ ID NOS: 4		
: SOFTWARE: FastSeq for Windows Version 4.0		
: SEQ ID NO 1		
: LENGTH: 3184		
: TYPE: DNA		
: ORGANISM: Homo sapiens		
: FEATURE:		
: NAME/KEY: CDS		
: LOCATION: (168)...(977)		
US-09-934-268-1		
Query Match	26.1%; Score 362; DB %;	Length 3184;
Best Local Similarity	66.9%; Pred. No. 6.8e-88;	
Matches 515; Conservative 0;	Mismatches 255; Indels 0;	Gaps
DG	95	AAGTAGCTGCCTGGTACAGTAAGTACTCCTTTCCAGCTACATCATATCTTCGTGGCTG 154   DB 202 AGTGCGGTGCTGCCGGAATAACTCTTCGTGGCTTCAAAATTGCTTTCGGGTGTGG 261   GY 155 GAGTTGTCCTCTTGGAGATCGGGCTGTGGGCATGGAGAGAAGGATGTGTCGACC 214   Db 262 GAGCCCCGTTCCTGACTATCGGCTCTTGAGCTTGGGCTGGAGAAAGGGCTTCTCGACA 321   GY 215 TCACCAAGTAGCCCGATGTCATGGAATCGAACCCCTGTGTGCTGCTGATGATGGTGGCG 274

Db	322	TTCTACGCCCTACAGACATCTGGAGAGCCCTTATACCCCGTGTGGCTGTTTGTGTGATGGAG	381
QY	275	TGTTGATGTTACCCCTGGGGGTTGGCCGGCTGCGTGGGGGCTCTGCGGAGAAATATCTGCT	334
Db	382	GCGTCATGTCGATGCTGTGGCTTTTCTGCTGTCATTTGGGGCCCTCCGGGAAACACCTTCC	441
QY	335	TGTCACACTTTTGTGGGACACATCTGGCGCATCTTCTTCGAGACTGGCTGTGGCCG	394
Db	442	TGCTCAAGTTTTTCTCCGTCTCTCGTCTGATCTTCTTCTGTGAGACTGGCAACAGGA	501
QY	395	TGCTGGCCCTCTCTGTCCAGGATGGGTGAGGAGCCGGTTCGGGACTTCTTGAGAGCA	454
Db	502	TCTGGGCTTTGTTCTTAAGGAGCTGGATTGGAGCCAGCTCAACCTTTCATCAACACA	561
QY	455	ACATCAGTCTCTACCGGGAGCATTTGCATGTGCACAAACCTATCGACTCCCTTCAGAAAG	514
Db	562	ACGTCAAGGCTTACCGGAGCACATTGACTCCGAACCTCATTTGACTTTGCTCAGGAAT	621
QY	515	CTAACCACTGCTGTGGCGCATATGGCCCTGAAGACTGGGACCTTCACGTCTACTTCATT	574
Db	622	ACTGGTCTTGTGTGAGAACCCGAGAGCCCAATGACTGGAACTCAATATCTACTTCANCT	681
QY	575	GCAGGCGTCCAGCTACAGCCGGAAGAAAGTCCGGGGTCCCTTCTCTGCGGTGCAG	634
Db	682	GCATGACTTGAACCCACGCGGGAGGCGCTCGGGGGTCCCTTCTCTGCTGCTCAGGG	741
QY	635	ATCCTGCGCAAAAAGTTGTGAACACACAGTGTGATATGATGTCAAGATTCAAGCTGAGA	694
Db	742	AACCTGGCGGAGGATGTCTTCAACACCCAGTGTGCTCTAGACCTCTCGGCTCAAACTGAGC	801
QY	695	GCAAGTGGGATGAGTCCATCTTCAACGAAAAGGCTGCATCCAGGGCGCTGGAAAAGCTGGCTC	754
Db	802	TGGAGCACACAGGGCTTCAATCAACACCAAAGGCTCGTGGGCCAATTGAGAAAGTGGCTGC	861
QY	755	CGCGGAACATTACATTGTGGCTGGCGTCTTCAATCGCATCTGCTGTGGCAATATTG	814
Db	862	AGGACAACCTGATGTGTGGCGGAGGATCTTCAATGGGCAATGCCCTCTCCACGATCTTGG	921
QY	815	GCATTTTCTTGGCAAGGACGTGATCTCAGACATTCGAGGCCATGTAAGCC	864
Db	922	GCATTTGCTTGGCCCGAAGCTCGTGAAGTCAATCAAGGCAAGTGAAGCC	971

```

      100          US-09-972-970-2
      100          ; Sequence 2, Application US/09972970
      100          ; Patent No. US20020164693A1
      100          ; GENERAL INFORMATION:
      100          ; APPLICANT: Shi et al.
      100          ; TITLE OF INVENTION: TM4SF Receptor Polynucleotides, Polypeptides, and Antibodies
      100          ; FILE REFERENCE: P0506P1
      100          ; CURRENT APPLICATION NUMBER: US/09/972,970
      100          ; CURRENT FILING DATE: 2001-10-10
      100          ; PRIOR APPLICATION NUMBER: PCT/US01/11130
      100          ; PRIOR FILING DATE: 2001-04-05
      100          ; PRIOR APPLICATION NUMBER: 60/195,336
      100          ; PRIOR FILING DATE: 2000-04-10
      100          ; NUMBER OF SEQ ID NOS: 8
      100          ; SOFTWARE: Patentin Ver. 2.0
      100          ; SEQ ID NO 2
      100          ; LENGTH: 2538
      100          ; TYPE: DNA
      100          ; ORGANISM: Homo sapiens
      100          US-09-972-970-2

Query Match       25.3%; Score 351.8; DB 9; Length 2538;
Best Local Similarity   66.9%; Pred. No. 3,4e-85;
Matches 517; Conservative    0; Mismatches 252; Indels 4; Gaps 1;

QY      92 CCAAGTCACGTGCGTGACAGTACTCCCTTTCACTCAACAATCATATCTCGGTGG 151
        ||| ||||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      150 CCGAGTCGGCGTGCGCGGAATCTTCCTTTGCCTTCACAAATGTTTCTGGGTGC 209

```

```

OY 152 CTGAGATGCTCTTCTTGGAGTGGGCTGTGGCATGAGCAAGGGTGTCTCG 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 TGGAGACCTGTCTCGGCGATCGGCTGTGGGCTGAGAGGGTGTCTCTCA 269
OY 212 ACCTACCAAGTGAACCGGATGATGATGATGATGATGATGATGATGATG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 ACATCTCTCGGCTGACGATCGGATGAGGCTGTGGGCTGTGGGCTGTGG 329
OY 272 GCGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 GAGGCGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389
OY 332 GCTTGTCAACCTTTTCTGTGGGATGATGATGATGATGATGATGATGAT 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 TCTGTCTCAAGTTTCTGATGATGATGATGATGATGATGATGATGATGAT 449
OY 392 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 GATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509
OY 452 GCAACATCAAGTCTTACCGGAGATGATGATGATGATGATGATGATGATG 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 ACAAGCTCAAGGCTCTATCGGAGATGATGATGATGATGATGATGATGATG 569
OY 512 AAGCTCAACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 AATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 629
OY 572 ATTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 ACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689
OY 632 CAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 690 GGGACCTGGG---ATGCTCTCAACACCCAGTGTGCTGCTGCTGCTGCT 745
OY 692 AGAGCAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATG 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 AGCTGAGGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 805
OY 752 TCCCGGCAACATTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 TGCAGGCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 865
OY 812 TTGGCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 TTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 918

```

```

RESULT 4
US-09-925-299-205
: Sequence 205, Application US/09925299
: Patent No. US20020055627A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA102
: CURRENT APPLICATION NUMBER: US/09/925, 299
: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05883
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124, 270
: NUMBER OF SEQ ID NOS: 1556
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 205
: LENGTH: 1655
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1548)

```

```

OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (1559)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (1564)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (1623)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (1643)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-205

Query Match      24.38; Score 337.4; DB 10; Length 1655;
Best Local Similarity 62.8%; Pred. No. 2,2e-81;
Matches 524; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

OY 33 GCGGCGGCGGCGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 GCTGCGGCGGCGAAGTGTCTTATGATGATGATGATGATGATGATGATGATG 302
OY 93 CAAAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 TGAAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
OY 153 TGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 GGAATTAACATTTTGTGAATGATGATGATGATGATGATGATGATGATGATG 422
OY 213 CCTCAACAAAGTCAAGCGGATGATGATGATGATGATGATGATGATGATGATG 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 CATCTCTTCAACACGATCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
OY 273 CGTGTGATGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 AGAGAGATGCTCAATTTGGAATTTGCAAGGCTGCTGCTGCTGCTGCTGCTG 542
OY 333 CTTCGCACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 CCTTCGCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
OY 393 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 AGTTCAGATTTGTTTCAAAAGCTGATGATGATGATGATGATGATGATGATG 662
OY 453 CAACATCAAGTCTACCGGAGATGATGATGATGATGATGATGATGATGATGAT 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 CAACATCAAGATATCGGAGATGATGATGATGATGATGATGATGATGATGATG 722
OY 513 AGCTAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 AATTTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
OY 573 TTGCAAGATTTCAATGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 TTGCAAGATTTCAATGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
OY 633 AGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 843 AGATCCCGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
OY 693 GAGCAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 AGTTGACACGACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 962
OY 753 CCGGCGGACATTTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 GCAAGCAATTTAACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
OY 813 TTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

1 TYPE: DNA
2 ORGANISM: Homo sapiens
3 FEATURE:
4 NAME/KEY: misc feature
5 LOCATION: (5)
6 OTHER INFORMATION: n equals a,t,g, or c
7 NAME/KEY: misc feature
8 LOCATION: (195)
9 OTHER INFORMATION: n equals a,t,g, or c
10 NAME/KEY: misc feature
11 LOCATION: (285)
12 OTHER INFORMATION: n equals a,t,g, or c
13 NAME/KEY: misc feature
14 LOCATION: (314)
15 OTHER INFORMATION: n equals a,t,g, or c
16 NAME/KEY: misc feature
17 LOCATION: (324)
18 OTHER INFORMATION: n equals a,t,g, or c
19 NAME/KEY: misc feature
20 LOCATION: (327)
21 OTHER INFORMATION: n equals a,t,g, or c
22 NAME/KEY: misc feature
23 LOCATION: (331)
24 OTHER INFORMATION: n equals a,t,g, or c
25 NAME/KEY: misc feature
26 LOCATION: (337)
27 OTHER INFORMATION: n equals a,t,g, or c
28 NAME/KEY: misc feature
29 LOCATION: (340)
30 OTHER INFORMATION: n equals a,t,g, or c
31 NAME/KEY: misc feature
32 LOCATION: (361)
33 OTHER INFORMATION: n equals a,t,g, or c
34 NAME/KEY: misc feature
35 LOCATION: (394)
36 OTHER INFORMATION: n equals a,t,g, or c
37 NAME/KEY: misc feature
38 LOCATION: (400)
39 OTHER INFORMATION: n equals a,t,g, or c
40 NAME/KEY: misc feature
41 LOCATION: (408)
42 OTHER INFORMATION: n equals a,t,g, or c
43 NAME/KEY: misc feature
44 LOCATION: (411)
45 OTHER INFORMATION: n equals a,t,g, or c
46 NAME/KEY: misc feature
47 LOCATION: (415)
48 OTHER INFORMATION: n equals a,t,g, or c

```

[illegible]

```

Db 374 TGAAGTCAGTGTGTCATCAAAATACCTCATATTTGGCTTCATATGTCATATTTGGTTT 433
QY 153 TGAAGTCAGTGTGTCATCAAAATACCTCATATTTGGCTTCATATGTCATATTTGGTTT 212
Db 434 GGGAAATACATTTCTTGGAATGAGCTGTGGCATGTGAAGAAAGAGGTTCTGTCCAA 493
QY 213 CCTCACAAAGTGAACCCGGATGATGAAATGACCCCTGTGTGCTGTGATGATGATG 272
Db 494 CATCTCTTCATCAACCCGATCTCGGCGCTTGAACCGATTGGCTCTTCTGTGTGTGG 553
QY 273 CGTGTGATGTTACCCGAGGTTGCGCGCTGCGTGGGGGCTCTCGGAGAAATATCG 332
Db 554 AGGAGTATGTTCAATTTGGGATTTGAGGTCATTTGAGGCGCTGAGGAAACACTTT 613
QY 333 CTTCCTCAACTTTTCTGTGTCACATCTGTCATCTTCTGTCGACCTGTGCTGCG 392
Db 614 CCTCTCAAGTTTCTGTGTCTGTCGAGAAATATTTCTTCTGACCTGACCTGCGG 673
QY 393 CGTGTGCTCTCTGTCACAGACTGTGAGGACCGCTTCCGGAGTCTTCTGAGAG 452
Db 674 ACTTCTACATTTGTTTTCAAAGACTGATCAACAGACCTGATTTCTTATAACAA 733
QY 453 CAACATCAAGTCTTACCGGAGCATGATCAATGCAAAACCTCATGACCTCCCTCAGA 512
Db 734 CAACATCAAGCATGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 793
QY 513 ACCTACCAAGTCTTACCGGAGCATGATGATGATGATGATGATGATGATGATGATG 572
Db 794 AATTTGAGGATGCTGTGGGCTTTTGGAGCTGATGATGATGATGATGATGATGATG 853
QY 573 TTGCAGCGAT 582
Db 854 TTGCACAGAT 863

```

```

RESULT 7
US-09-729-674-169
; Sequence 169, Application US/09729674
; Patent No. US2001003935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 169
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-169

```

```

Query Match 16.5%; Score 229; DB 10; Length 1110;
Best Local Similarity 57.1%; Pred. No. 3,1e-52;
Matches 477; Conservative 2; Mismatches 292; Indels 64; Gaps 1;
QY 33 GCCCGCCGCCGCCGCGGATCTGCTTCTCAGAGATGCACTATTTATAGATCTCTACGC 92

```

```

Db 142 GCTGCCGGGGGCAAGTGTCTTCTATGAAACAGAGATGTCGGGGAAGCACTACAGGGTCC 201
QY 93 CAAGTCACTCTGTGACAGTACCTCTTTTACGTACACATCATCTTCTGTGTGC 152
Db 202 TGAAGTCAGTGTGTCATCAAAATACCTCATATTTGGCTTCATATGTCATATTTGGTTT 261
QY 153 TGAAGTGTCTTCTGTGAGTGGGCTGTGGGATGAGGAAAGAGTGTGCTGCCA 212
Db 262 GGGAAATACATTTCTTGAAATTTGACCTGTGGCATTTGATGAAAGAAAGATTTGTGCCA 321
QY 213 CCTCACAAAGTGAACCCGGATGATGAAATGACCCCTGTGTGCTGTGATGATGATG 272
Db 322 CATCTCTTCATCAACCCGATCTCGGCGCTTGAACCGATTGGCTCTTCTGTGTGTGG 381
QY 273 CGTGTGATGTTACCCGAGGTTGCGCGCTGCGTGGGGGCTCTGCGGAGAAATATCG 332
Db 382 AGGAGTATGTTCAATTTGGGATTTGAGGATTTGAGGATTTGAGGATTTGAGGATTTG 441
QY 333 CTTCCTCAACTTTTCTGTGTCACATCTGTCATCTTCTGTCGACCTGTGCTGCG 392
Db 442 CTTCCTCAAGTTTCTGTGTCTGTCGAGAAATATTTCTTCTGAGCTGACCTCGG 501
QY 393 CGTGTGCTCTCTGTCACAGACTGTGAGGACCGCTTCCGGAGTCTTCTGAGAG 452
Db 502 AGTTCTAGCATTTGTTTCAAGACTGATCAAGACAGCTGATTTCTTTTAAACAA 561
QY 453 CAACATCAAGTCTTACCGGAGCATGATGATGATGATGATGATGATGATGATGATG 512
Db 562 CAACATCAAGCATGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 621
QY 513 AGCTAACCAAGTGTGTGGGCAATGAGGCTGAGAGCTCAACGTCATCTTCAA 572
Db 622 ATATA----- 626
QY 573 TTGCAGCGGTCAGTACAGCCGAGAGAGTGGGGTCCCTCTCTCTGCTGCTGCC 632
Db 627 -----TTCCAAATGCAAGTGCAGAGCATGTGGCGCTTCATCTCTGCTGCACTAA 677
QY 633 AGATCCTGGCAAAAGTGTGACACACAGTGTGATGATGATGATGATGATGATGATG 692
Db 678 AGATCCCGAGAGATGTATCAACATGATGATGATGATGATGATGATGATGATGATG 737
QY 693 GAGCAAGTGGATGAGTCCATCTTCACGAAAGGCTGATCCAGGCGCTGGAAGCTGGCT 752
Db 738 AGTTGACCAAGATGTATCTACAGAAAGGCTGTGTGCCCATTTGAGAAAGTGTT 797
QY 753 CCGCGGAACTTTACATGTGTGGTGGCTGTATGCTCATGCTGCTGTGTGTGATGAT 812
Db 798 GCAGGACAAATTTACCCWCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 857
QY 813 TGGCATCTCTCTGCGCAAGGACGCTGATCTCAGACATGAGGACAGTGAAGCGCGC 867
Db 858 TGGGATMTGCTGTGGCCAGAAATTTGGTATGAGATTCGATGCTGAGGCGAGC 912

```

```

RESULT 8
US-09-764-873-12
; Sequence 12, Application US/09764873
; Patent No. US20020151479A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7201
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA REMOVED - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 12
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-09-764-873-12

Query Match 11.9%: Score 165.2; DB 10; Length 545;  
 Best Local Similarity 98.28; Pred. No. 3.6e-35;  
 Matches 167; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CTCGGCCGACCGCGCGCGCGCGCTGCCCGCGCGCGCGGATTCCTCTCAGA 64  
 |||  
 Db 9 CTCGGCCGACCGCGCGCGCGCGCTGCCCGCGCGCGCGGATTCCTCTCAGA 68  
 |||  
 Qy 65 AGATGCACTATTAGTACTCTAACGCAAGTACGCGCTGACAGTACCTCTT 124  
 |||  
 Db 69 AGATGCACTATTAGTACTCTAACGCAAGTACGCGCTGACAGTACCTCTT 128  
 |||  
 Qy 125 TCACCTACCAACATCATCTTCTGCTGGCTGAGTTGCTCTTGGAGTC 174  
 |||  
 Db 129 TCACCTACCAACATCATCTTCTGCTGGCTGAGTTGCTCTTGGAGTC 178  
 |||

RESULT 9

US-10-052-586-277  
 Sequence 277, Application US/10052586  
 Patent No. US20020127584A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Chen, Jian  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Pan, James  
 APPLICANT: Smith, Victoria  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C1

CURRENT APPLICATION NUMBER: US/10/052.586

PRIOR FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059266

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/063120

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063121

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063486

PRIOR FILING DATE: 1997-10-21

PRIOR APPLICATION NUMBER: 60/063540

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063541

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063564

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063734

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/063870

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/064103

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066120

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: 60/066466

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/066772

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/069335

PRIOR FILING DATE: 1997-12-11

PRIOR APPLICATION NUMBER: 60/069425

PRIOR FILING DATE: 1997-12-12

PRIOR APPLICATION NUMBER: 60/069870

PRIOR FILING DATE: 1997-12-17

PRIOR APPLICATION NUMBER: 60/068017

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/07450

PRIOR FILING DATE: 1998-03-10

PRIOR APPLICATION NUMBER: 60/077632

PRIOR FILING DATE: 1998-03-11

PRIOR APPLICATION NUMBER: 60/077649

PRIOR FILING DATE: 1998-03-11

PRIOR APPLICATION NUMBER: 60/078886

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/078939

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/079664

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079786

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/080107

PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080194

PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080327

PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/080333

PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/081049

PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/081070

PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/081195

PRIOR FILING DATE: 1998-04-09

PRIOR APPLICATION NUMBER: 60/081838

PRIOR FILING DATE: 1998-04-15

PRIOR APPLICATION NUMBER: 60/082568

PRIOR FILING DATE: 1998-04-21

PRIOR APPLICATION NUMBER: 60/082569

PRIOR FILING DATE: 1998-04-21

PRIOR APPLICATION NUMBER: 60/082704

PRIOR FILING DATE: 1998-04-22

PRIOR APPLICATION NUMBER: 60/082797

PRIOR FILING DATE: 1998-04-22

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/083495

PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083496

PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083499

PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083559

PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/084366

PRIOR FILING DATE: 1998-05-05

PRIOR APPLICATION NUMBER: 60/084414

PRIOR FILING DATE: 1998-05-06

PRIOR APPLICATION NUMBER: 60/084639

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084640

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084643

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/085573

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/085579

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/085580

PRIOR FILING DATE: 1998-05-15



```

; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
US-10-042-417-37

```

Query Match	8.28;	Score 113.4;	DB 12;	Length 368;
Best Local Similarity	65.88;	Pred. No. 2.8e-21;		
Matches 194;	Conservative 0;	Mismatches 98;	Indels 3;	Gaps 2;

[illegible]

```

RESULT 11
US-09-764-873-14
: Sequence 14, Application US/09764873
: Patent No. US20020151479A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT001
: CURRENT APPLICATION NUMBER: US/09/764,873
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PAMM or file wrapper
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 14
: LENGTH: 288
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-764-873-14

```

[illegible]

RESULT 12  
 US-09-960-352-1269  
 : Sequence 1269, Application US/09960352  
 : Patent No. US20020137139A1  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Warren, Wesley C.  
 : APPLICANT: Tao, Nengping  
 : APPLICANT: Byatt, John C.  
 : APPLICANT: Mathilagan, Nagappan  
 :  
 : TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 : FILE OF INVENTION: MUSCLE AND FAT DEPOSITION  
 : FILE REFERENCE: 16511.006/37-21(10298)C

```

?
? CURRENT APPLICATION NUMBER: US/09/960,352
? CURRENT FILING DATE: 2001-09-24
? NUMBER OF SEQ ID NOS: 15112
? SEQ ID NO 1269
? LENGTH: 441
?
? TYPE: DNA
? ORGANISM: Bos taurus
? OTHER INFORMATION: Clone ID: 06-LIB3058-036-Q1-K1-B5
? US-09-960-352-1269

```

Query Match	5.28;	Score 72;	DB 10;	Length 441;
Best Local Similarity	51.68;	Pred. No. 44e-10;		
Matches 190;	Conservative	0;	Mismatches 175;	Indels 3; Gaps
QY	97	GTCACTGCTGGTACAACTACCTCCTTTTCAGCTACACATCATCTTGTGGCTGGA	156	
DB	30	GGCACCAAGTGCATCAATATACCTGCTCTGGATTTAACTTCATCTTGTGGCTGGCGG	89	
QY	157	GTGTCTCTTTGGAGTGGGCTGTGGCATGAGACCAAAAGGTGTGCTGTCCGACTC	216	
DB	90	ATCCGGGCTCTTCTGTGGAGCTATGGCTCCGATTCGACTCGACCAAGAGCATCTTC	149	
QY	217	ACCAAGTGACCCGGATGCATGATGACGACCCCTGTGGTGTGCTGTGATGTGGCGTG	276	
DB	150	GAGCAAGAAATTAATGATATTCACGCTTCTACACAGCGCTATATCTTGATCGGAGCCGT	209	
QY	277	GTGATGTTCAACCCTGGG---TTGGCCGGCTGCGTGGGGGCTTCGCGGAGATATCTGC	333	
DB	210	GCGCTCATGATGCTGGTGGGCTCTCTGGGCTGTGCGGAGCGCGTGGAGGTCCACTGC	269	
QY	334	TTGGTCACCTTTTCTGTGGGACCATGCTGCTCATCTTCTCCGGAAGTGGCTGTGGCC	393	
DB	270	ATGCTGGGATGTTGTTTACCTTCTCTTGGTGATATTTGGCATTTGAAGTAGCTGCAGCC	329	
QY	394	GTGCTGGGCTTCGTGTTCCAGAGACTGGGTGAAGGACCGGTTCCGGAGTCTTTCAGAGC	453	
DB	330	ATCTGGGATATATCCACCAAGAGAGGTGATCAAGGAAGTCCAGAGATTTTACGAGGAC	389	
QY	454	AAATCAAA 461		
DB	390	ACCTACAA 397		

```

RESULT 13
US-09-895-828-318
; Sequence 318, Application US/09895828
; Patent No. US20020099012A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Carter, Darrick
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.539
; CURRENT APPLICATION NUMBER: US/09/895,828
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 318
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4, 5
; OTHER INFORMATION: n = A,T,C or G
US-09-895-828-318

Query Match          5.0%; Score 69.8; DB 10; Length 547;
Best Local Similarity 53.1%; Pred. No. 1.9e-09;

```





QY 447 CGAGAGCAACATCAA 461  
| | | | |  
DB 475 CAAGGACACCTTACAA 489

Search completed: December 19, 2002, 06:22:12  
Job time : 177 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 19, 2002, 05:20:37 ; Search time 2589 seconds

(without alignments)  
3035.055 Million cell updates/sec

Title: US-09-905-674-2

Perfect score: 1449

Sequence: 1 MHYRYRNAKSCWYKYL.....IFLARLISDIENAKAGHHF 270

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO/US0905674/runat\_16122002\_111845\_6756/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastlap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEPSTLZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0905674.ecgn\_1.1.1616\_erunat\_16122002\_111845\_6756 -ICPU=3  
-NO\_XLPHY -NO\_MAP -LANG=EUTRY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vl:\*  
30: em.htg.hum:\*  
31: em.htg.inv:\*  
32: em.htg.other:\*  
33: em.htg.mus:\*  
34: em.htg.pin:\*  
35: em.htg.rod:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.sy:\*  
39: em.htgo.hum:\*  
40: em.htgo.mus:\*  
41: em.htgo.other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1449	100.0	1388	6 AX440923	AX440923 Sequence
2	1449	100.0	2553	9 AF311903	AF311903 Homo sapi
3	1438	99.2	2556	9 HSM601611	AL136638 Homo sapi
4	1420	98.0	2428	10 BC025568	BC025568 Mus muscu
5	1420	98.0	2498	10 BC024611	BC024611 Mus muscu
6	1420	98.0	2500	10 BC026574	BC026574 Mus muscu
7	1338.5	92.4	2502	9 BC002920	BC002920 Homo sapi
8	867	59.8	1405	9 AF065389	AF065389 Homo sapi
9	867	59.8	1416	9 BC009704	BC009704 Homo sapi
10	867	59.8	3175	10 AF121344	AF121344 Mus muscu
11	861	59.4	1516	10 BC010346	BC010346 Mus muscu
12	858	59.2	813	6 AX420468	AX420468 Sequence
13	858	59.2	3184	6 AX420466	AX420466 Sequence
14	849	58.6	1408	9 AF053455	AF053455 Homo sapi
15	825	56.9	2426	9 BC010405	BC010405 Homo sapi
16	572	39.5	1995	10 BC024685	BC024685 Mus muscu
17	568	39.2	852	6 AX247836	AX247836 Sequence
18	567	39.1	864	6 AX343015	AX343015 Sequence
19	559	38.6	1567	6 AX061778	AX061778 Sequence
20	522.5	36.1	102468	2 AC009984	AC009984 Drosophila
21	522.5	36.1	106584	2 AC013977	AC013977 Drosophila
22	522.5	36.1	140973	3 AC008140	AC008140 Drosophila
23	522.5	36.1	159007	3 AC009219	AC009219 Drosophila
24	522.5	36.1	224795	3 AE003688	AE003688 Drosophila
25	515	35.5	4445	9 AK024427	AK024427 Homo sapi
26	508	35.1	140468	2 AC017377	AC017377 Drosophila
27	508	35.1	174832	3 AC0092216	AC0092216 Drosophila
28	508	35.1	174832	3 AC0092216	AC0092216 Drosophila
29	508	35.1	260550	3 AE003612	AE003612 Drosophila
30	503	34.7	1216	3 AF274013	AF274013 Drosophila
31	451.5	31.2	22920	3 CERI4G10	268880 Caenorhabditis
32	433	29.9	11172	2 AC013030	AC013030 Drosophila
33	433	29.9	191342	3 AC104140	AC104140 Drosophila
34	433	29.9	289090	3 AE003424	AE003424 Drosophila
35	406	28.0	56820	3 DMBRC10	AF121804 Drosophila
36	404.5	27.9	1703	9 AF120266	AF120266 Homo sapi
37	404.5	27.9	1726	9 BC003157	BC003157 Homo sapi
38	404.5	27.9	1734	9 BC004161	BC004161 Homo sapi
39	395.5	27.3	1778	6 AX376210	AX376210 Sequence
40	393	27.1	1661	10 BC003872	BC003872 Mus muscu
41	385	26.6	1646	10 BC019991	BC019991 Mus muscu
42	370	25.5	1601	9 AF325213	AF325213 Homo sapi
43	370	25.5	1693	9 BC032802	BC032802 Homo sapi
44	349	24.1	557	6 AX079573	AX079573 Sequence
45	343.5	23.7	860	10 RNAD1A	X61654 Rat mRNA fo

RESULT 1

#### ALIGNMENTS

AX440923 1388 bp DNA linear PAT 28-JUN-2002  
LOCUS AX440923  
DEFINITION Sequence 1 from Patent WO0206340.  
ACCESSION AX440923  
VERSION AX440923.1 GI:21665563  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 Reinhard, C.J. and Garcia, P.D.  
Tetraspan protein and uses thereof  
Patent: WO 0206340-A 1 24-JAN-2002;  
CHIRON CORPORATION (US)  
FEATURES  
source 1.1388  
location/Qualifiers  
BASE COUNT 270 a 390 c 411 g 313 t 4 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,37e-154 Length: 1388  
Score: 1449.00 Matches: 270  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-905-674-2 (1-270) x AX440923 (1-1388)  
QY 1 MethISTYTYRATGTYRYSerASnaAlaLysValSerCysTRPtyrLysTYrLeuPhe 20  
|||||  
Db 67 ATGACATATATATAGATCTTACGCCAAGTCAGCTGGTACAAAGTACCTCTTTC 126  
QY 21 SerTYrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyAlaGlyLeuTrpAla 40  
|||||  
Db 127 AGCTACAAACATCACTTCTGGTGGCTGGAGTTGCTTCTTGGAGTCCGGCTGGGGCA 186  
QY 41 TrpSerGluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60  
|||||  
Db 187 TGGAGGAAAGGCTGTGCTGCTCCGACCTCACCAGTACCCGGATGATGATCGAC 246  
QY 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80  
|||||  
Db 247 CCTGTGCTGCTGCTGCTGATGGTGGCGGCTGGATGTCACCTGGCGCTTCCCGGCTGC 306  
QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100  
|||||  
Db 307 GTGGGGGCTGCGGGAGAAATATCTCTTCTCACTTTTCTGTGGCCATCTGCTGC 366  
QY 101 IlePhePheLeuGluLeuAlaValAlaValAlaPheLeuPheGlnAspTrpValArg 120  
|||||  
Db 367 ATCTTCTTCTGAGCTGGGCTGTGGCGCTGGCTTCTCTGTCAGAGACGTGGGTGAGG 426  
QY 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTYrArgAspAspIleAspLeu 140  
|||||  
Db 427 GACCGGTTCCGGGAGTTCTTCGAGCAACATCAAGTCTCCTACGGGAGCATATCATCTG 486  
QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyValAlaThrGlyProGlu 160  
|||||  
Db 487 CAATACCTATGCTGCTGCTGCTGAGAAAGCTAACCAAGTGTGGGCAATATGCCCTGAA 546  
QY 161 AspTrpAspLeuAsnValTYrPheAsnCysSerGlyAlaSerTYrSerArgLysCys 180  
|||||  
Db 547 GACTGGGACCTCAAGCTTACTTCAATTCAGCGGTGCGAGCTACAGCCGAGAGAGAGTC 606  
QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCys 200  
|||||  
Db 607 GGGGTCCTCTTCTGCTGCTGCGGAGATCCCTGCGCAAAAGTGTGACACACAGTGT 666  
QY 201 GlyTYrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220

|||||  
Db 667 GGATATGATGTCAGGATTCAGTGAAGAGCAAGTGGATGATGTCATCTTCACGAAAGGC 726  
QY 221 CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTYrIleValAlaGlyValPhe 240  
|||||  
Db 727 TGCATCCAGGGCGCTGGAAGCTGGCTCCCGGGAACATTTACATTTGGCTGGCGTCTTC 786  
QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260  
|||||  
Db 787 ATCGGCATCTGCTGTTGTCAGATATTTGGCATCTTCTCTGGCAAGACGCTATCTCGAC 846  
QY 261 IleGluAlaValLysAlaGlyHisHisPhe 270  
|||||  
Db 847 ATCGAGCGAGTGAAGCCGCGCATCATCTTC 876  
RESUIT 2  
AF311903  
LOCUS AF311903 2553 bp mRNA linear PRI 16-AUG-2002  
DEFINITION Homo sapiens DC-TM4F2 precursor, mRNA, complete cds.  
ACCESSION AF311903  
VERSION AF311903.1 GI:2266721  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 2553)  
AUTHORS Zhang, W., Li, N., Wan, T., and Cao, X.  
TITLE Identification of novel membrane proteins  
JOURNAL Unpublished  
2 (bases 1 to 2553)  
REFERENCE  
AUTHORS Zhang, W., Li, N., Wan, T., and Cao, X.  
TITLE Direct Submission  
JOURNAL Submitted (06-OCT-2000) Department of Immunology, Second Military  
Medical University & Shanghai Brilliance Biotechnology Institute,  
800 Xiangyin Rd., Shanghai 200453, P.R.China  
FEATURES  
source 1.2553  
location/Qualifiers  
CDS  
1.2553  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
73.885  
/codon\_start=1  
/product="DC-TM4F2 precursor"  
/protein\_id="AA094899.1"  
/db\_xref="GI:2266722"  
/translation="MHYRYSNAKVSQWRYLFSYNIIFMLAGVFLGLWAMSEK  
GVLSDLTKYVRMGIDIPVYLIMGVMTLIGAGCGVALREUGLINFRCGYILIF  
FLDLAVAVLAFLEQDNYRDRREFEFESNFKSRDDIDLNLDSLQANQCCGATPE  
DMDLNYFNGSGASYSREKGVFSCVDPKRVVNTCCGIDVRIQLSKWDESLFT  
KGIQALLESWLPNNIYIVAGVFIAISLQIFIFLARTLISDIEAVKAGHP"  
BASE COUNT 505 a 724 c 680 g 644 t  
ORIGIN  
Alignment Scores:  
Pred. NO.: 1.4e-153 Length: 2553  
Score: 1449.00 Matches: 270  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-905-674-2 (1-270) x AF311903 (1-2553)  
QY 1 MethISTYTYRATGTYRYSerASnaAlaLysValSerCysTRPtyrLysTYrLeuPhe 20  
|||||  
Db 73 ATGACATATATATAGATCTTACGCCAAGTCAAGCTGGTACAAAGTACCTCTTTC 132  
QY 21 SerTYrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyAlaGlyLeuTrpAla 40  
|||||  
Db 133 AGCTACAAACATCACTTCTGGTGGCTGGAGTTGCTTCTTGGAGTCCGGCTGGGGCA 192  
QY 41 TrpSerGluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60  
|||||

Db	193	TGGAGCGAAAAGGGGTGCTCTGTCCGACCTCACCAAGGACCCGGATGATCATGAAATGCAC	252
OY	61	ProValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys	80
Db	253	CGTGGGGTGCTGTGCTCTGATGTGGGGCTGTGTGATGTTCACCCGTGGGGTGTGGCGGGCTGC	312
OY	81	ValGlyAlaLeuValGlyLeuAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu	100
Db	313	GTGGGGGCTGTGCGGGGAATATCTGCTTGGTCAACTTTTCTGTGGCACCATGTGTGCTC	372
OY	101	IlePhePheLeuGlyLeuAlaValAlaValLeuAlaPheLeuPheGlnAsnTyrValArg	120
Db	373	ATCTTCTTCTCGAGACCTGGCTGTGGCCCTGCTGTGGCTTCTCTGTTCACGAGACTGGGTGAGG	432
OY	121	AspArgPheArgGlyPhePheGlyLeuSerAsnIleLeuSerTyrArgAspAspIleAspLeu	140
Db	433	GACCGGTTCGGGAGACTTCTCGAGAGCAACATCAAGTCTCACGGGAGCATATGATGTG	492
OY	141	GlnAsnLeuIleAspSerLeuGlnIleValAlaAsnGlnCysGlyAlaTyrGlyProGlu	160
Db	493	CAAAACCTCATCGACTCCCTTCAGAAACCTAACACAGTCTGTGGCGCATATGGCCCTGAA	552
OY	161	AspTyrPaspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlnLysCys	180
Db	553	GACTGTGGACCTCACAGTCTACTTCAATTGCACGGGTGCCACTCACAGCCGAGAGAATGC	612
OY	181	GlyValProPheSerCysCysValProAspProAlaGlnValValAsnThrGlnCys	200
Db	613	GGGGTCCCTTCTCTCTGTGGTGGTGCACAGTCTGTGGCAAAAAGTTGTGACACACAGTGT	672
OY	201	GlyTyrAspValArgIleGlnLeuLysSerTyrTyrAspGlySerIlePheThrLysGly	220
Db	673	GGATGTGATGTGCAGGATTACACTGAAAGCAAGTGGATGTGATTCATCTTCACGAAAGGC	732
OY	221	CysIleGlnAlaLeuGlnSerTyrLeuProArgAsnIleTyrIleValAlaGlyValPhe	240
Db	733	TGCATCCAGCGGCTGTGGAAGCTGGCTCCCGGGAACATTTTCATTGTGGCTGGGCTCTC	792
OY	241	IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp	260
Db	793	ATCGGCATCTGCTGTGTGCAGATATTGGCATCTTCCGTGGCAGAGCCGTATCTCAAC	852
OY	261	IleGluAlaValLysAlaGlyHisHisPhe	270
Db	853	ATCGAGCGCATGAAGCGCGGCATCATCTTC	882
RESULT 3			
HSM801611			
LOCUS	HSM801611	2556 bp	mRNA
DEFINITION	Homo sapiens mRNA; CDNA DKFP2564B1037 (from clone DKFP2564B1037); complete cds.		linear PRI 12-JUL-2002
ACCESSION	AL136638		
VERSION	ALI36638.1		
KEYWORDS	GI:12052801		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 2556)		
TITLE	Pousta,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.		
JOURNAL	Direct Submissiion		
COMMENT	Submitted (10-JUL-2002) GSF, Institut fuer Bioinformatik MIPS, Ingolstaedter Landstr. 1, 85764 Neuberg, Germany Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de; Sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFP2564B1037) is available at the RPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <a href="http://mips.gsf.de/proj/cDNA/">http://mips.gsf.de/proj/cDNA/</a> .		

FEATURES		Location/Qualifiers
source		1. .2536 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="10" /clone="DKR2p564B1037" /tissue_type="brain" /clone_id="564 (synonym: hfbr2). Vector pAMP1; host X1-2blue; sites NotI + SalI" /dev_stage="fetal" 1. .2556 /gene="DKR2p564B1037" 79. .891 /gene="DKR2p564B1037" /note="strong similarity to tetraspan NER-4" /codon_start=1 /product="hypothetical protein" /protein_id="CA666573.1" /db_xref="GI:12052802" /db_xref="SPRMBL:Q9H001" /translation="MHYRYSNAKVCWMYKYLFSYNIIFWLAGVFLGYOLAMSERK GLDLIRTRPMHGIDPVLAVGVVMEFLFGVCGALBENCLNPFRCGTVLIEF FLEIADVAVLAFIPQDWDWRDFREFEENSINISRDIDLVNIDLSLOKANOCCGAYGE DMDLVAVFNFCGASYSNRKCGVFPSCCYPRQAQVYVNTQCCYDPIRLOKSMDSEIFTF KGCIOALESWLPNRIYVAGVFAISLDLIFGIFLARTLSDIEAVRAGRPF" 2546 polya_site BASE COUNT 500 a 729 c 683 g 644 t ORIGIN
Alignment Scores:		
Pred. No.:	2.45e-152	Length: 2556
Percent:	1438.00	Matches: 268
Percent Similarity:	99.63%	Conservative: 1
Best Local Similarity:	99.26%	Mismatches: 1
Query Match:	99.24%	Indels: 0
DB:	9	Gaps: 0
US-09-905-674-2 (1-270) x HSM801611 (1-2556)		
QY	1	MethISTYRTYArgTYSerASnaAlaYSValISerCYSTRPYTLySTytleuEuphe 20
Db	79	ATGCACCTATTAAAGTACTTAAAGCCCAAGGACACTCTGGTAAAGTACCTCCTTTC 138
QY	21	SerTYrAsnIleIlePheTrPLeuNlaGlyValAlPheLeuGlyValGlyLeuTrpAla 40
Db	139	AGCTCAACATATCTTCTGCTGGCTGGAGTGTCTTCTCTTGGAGTGGGGCTGGGCA 198
QY	41	TrpSerGIuLysGlyValLeuSerAspLeuThrIysValThrArgMetHisGlyIleAsp 60
Db	199	TGGAGCGAAAGAGGTGGTGTGCCAGCCACCAAAAGTGACCCGGATGATGAAATCGAC 258
QY	61	ProValValLeuValLeuMetValGlyValAlMetPheThrLeuGlyPheAlaIysCys 80
Db	259	CCTGTGGTGGTGGTCTGTGTGGGCGGTGAGATGTCACCCCTGGGGTTCGCCGGCTGC 318
QY	81	ValGIValAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
Db	319	GTTGGGGGCTCTCGGGAGATATCTGCTTGTCTTCAACTTTTCTGTGGGACCATCTGCTC 378
QY	101	IlePhePheLeuGluIleuAlaValAlaValLeuAlaPheLeuPheGluAspTrpValArg 120
Db	379	ACCTCTTCCTCGAGACTGGCTGTGGCCGTCGTCCTTCCTTCAGAGACTGGGTGAGG 438
QY	121	AspArgPheArgGluPhePheGluSerAsnIleIysSerTYrArgAspAspIleAspLeu 140
Db	439	GACCGGTTCCGGGAGATTCTTCGAGAGCAACAACAGTCCCTACCGGAGATATCATCTG 498
QY	141	GluAsnLeuIleAspSerLeuGluIysAlaAsnGlnCysGlyAlaIleTyrlGlyProGlu 160
Db	499	CAAAACGTCATGATCTCCCTTACGAAAGCTAACCGAGTCTGTGGGGCATATGGCCCTGAA 558
QY	161	AspTrpAspLeuAsnValIlyrPheAsnCysSerGIyAlaSerTYrSerArgIuLysCys 180

|||||  
Db 559 GACTGGACCTCAACGCTACTCAATGACGGGTGCAGCTACAGCCGAGAGAGATGC 618  
QY 181 G1YVALPropheserCysValProAspProAlaGlnValValAsnThrGlnCys 200  
Db 619 GGGGTCCCTTCCTCCGCTGCGTGCAGATCCCTGCCAAAAGTTGTGAACACAGCTGT 678  
QY 201 G1TYrAspValArgIleGlnLeuLysSerLysTrpAspGlnSerIlePheThrLysGly 220  
Db 679 GGAATGATGATCAGATTGACGTGAAGACCACTGGGATGATGATCTTCACGAAAGGC 738  
QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyAlaPhe 240  
Db 739 TGCATCAGAGCCCTGGAAGAGCTGCTCCGCGAACAATTTACATGTGCGCTGCTTC 798  
QY 241 ILeAlaIleSerLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260  
Db 799 ATGCCATCTCCGCTGTGGCAGATATTGGCATCTTCCTGCAAGACGCTGATCTCAGAC 858  
QY 261 ILeG1AlaValAlaGlyAlaGlyHisIlePhe 270  
Db 859 ATCGAGCAGATGAAGCCGCGCTCCTCCTTC 888  
RESULT 4  
BC025568 2428 bp mRNA linear ROD 07-AUG-2002  
LOCUS Mus musculus, similar to region containing hypothetical protein  
DEFINITION MGCL1352; slingshot, clone MGC:36595 IMAGE:5322531, mRNA, complete  
cds:  
ACCESSION BC025568  
VERSION BC025568.1 GI:19343883  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2428)  
REFERENCE Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-7590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNU)  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgcen@rl.nih.gov](mailto:nisc_mgcen@rl.nih.gov)  
Blakesley, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, U., Pearson, R., Stanlipop, S., Thomas, P.J., Touchman, J.W.,  
Tsuang, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNU at: <http://Image.llnl.gov>  
Series: IRAN Plate: 56 Row: J Column: 21  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Genomescan gene  
prediction.  
FEATURES  
Source Location/Qualifiers  
1..2428  
/organism="Mus musculus"  
/db\_xref="taxon:10090"

/map="FVB/N"  
/clone="MGC:36595 IMAGE:5322531"  
/tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating  
ductal carcinoma, 5 month old virgin mouse."  
/clone\_id="NCI\_CGAP\_Mam6"  
/lab\_host="DH10B"  
/note="Vector: pCMV-Sport6"  
64..876  
/codon\_start=1  
/product="Similar to region containing hypothetical  
protein MGCL1352; slingshot"  
/protein\_id="AAH25568.1"  
/db\_xref="GI:19343884"  
/translation="MHYRYSNAEVSQWRYLFLFSYIVEMLAGVVELDGLWAMSEK  
GVSLDKTRRLGIDPVYILVMGVMEIFRAGCGVALREMICLRFCCAYILIF  
FLDLAVAVLAFLEQDWRDRFRFEFESNLSKSRDDLDLNLDSLOKAOCCGATPE  
DWDLNVFNCGSASYSERKGVPSCCVPDPAPKVVNTCCGIDVRLOLSKDFLETF  
KGCIOALEGWLPRNIYIVAGVFVIAISLQIFGLFRLTISDIEAVKAGHHF"  
BASE COUNT 517 a 640 c 624 g 647 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,48e-150 Length: 2428  
Score: 1420.00 Matches: 263  
Percent Similarity: 98.52% Conservative: 3  
Best Local Similarity: 97.41% Mismatches: 4  
Query Match: 98.00% Indels: 0  
DB: 10 Gaps: 0  
US-09-905-674-2 (1-270) x BC025568 (1-2428)  
QY 1 MethIstYrYrArgTyrSerAsnAlaLysValSerCysTrpIlyLysTyrLeuPhe 20  
Db 64 ATGCACCTTTATGATACTCGACGCCGAGGTGACGTCTGATCAAGTACTCTCTTC 123  
QY 21 SerTyrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAla 40  
Db 124 AGCTACATATCGCTCTTTGGCTGGCGAGATGTCTCTCTTGAGTGGCGGCTGGGCA 183  
QY 41 TrpSerGlyLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60  
Db 184 TGGAGCGAAAGGGGTGCTGCTCCGACCTCACCAAGGTGACCCGCTGATGATGAC 243  
QY 61 ProValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys 80  
Db 244 CCCGCGCGCTGCTTGATGATGCTGGCGCTGTGATGTTACACACGCGGATTCGACGCTGT 303  
QY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100  
Db 304 GTGGGGGCGCTCCGAGAGAACATCGCTGCTCAAGTTTCTGTGGGCGCATTTGCTCTC 363  
QY 101 IlePhePheLeuGlnLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg 120  
Db 364 ATCTTCTTCCTGGAACGTGGCGCTGCGCTGTGGCTTTTATVCCAAAGACTGGGTAGA 423  
QY 121 AspArgPheArgGlyPhePheGlnSerAsnIleLysSerTyrArgAspAspIleAspLeu 140  
Db 424 GACCGGTTCCGGGAATTCCTTGAGACCAACATCAAGTCTATGCGGATGACATCGACTG 483  
QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGln 160  
Db 484 CAGAACCTCATTTGACTCCCTTCAGAAAGCTAATCAGTGTGCGGGGCTTCGCGCCCTGAA 543  
QY 161 AspTrpAspLeuAsnValTyrPheAsnGlnSerLysValAserTyrSerArgGlyLysCys 180  
Db 544 GACTGGGACCTCAATGCTACTTCACTGCACTGCTCCAGCTACACCCAGGAAATGT 603  
QY 181 G1YVALPropheserCysValProAspProAlaGlnValValAsnThrGlnCys 200  
Db 604 GGGGTACCTTCTCTGCTGTGTGCGCAGATCTGCAAAAAGTCTGGAACACACAGTGT 663  
QY 201 G1TYrAspValArgIleGlnLeuLysSerLysTrpAspGlnSerIlePheThrLysGly 220  
|||||

Db 664 GGCTATGATGTCGGGATTCAGTGAAGACCAAGGGATGAGTTCATCTTACAAAAGCA 723

QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyAlaPhe 240  
|||||  
Db 724 TGCATCCAGGCTCTGGAAAGCGTGGCTGCCAGAACATCTACATGTGTGCTGTCTTC 783

QY 241 ILeAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260  
|||||  
Db 784 ATGGCATCTCAGCTGTCAGATTTTGGCATCTCTCGAGAGACCTGATCTCAGAC 843

QY 261 ILeGlnAlaValIleAlaGlyHisHisPhe 270  
|||||  
Db 844 ATCGAGGACGATGAAGGACGACATCATCTTC 873

RESULT 5  
BC024611

LOCUS BC024611 2498 bp mRNA linear ROD 07-AUG-2002

DEFINITION Mus musculus, similar to region containing hypothetical protein MGCI1352; slingshot, clone MGC:28503 IMAGE:4188261, mRNA, complete cds.

ACCESSION BC024611  
VERSION BC024611.1 GI:19354053

KEYWORDS MGC.

SOURCE house mouse.

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2498)

REFERENCE Strausberg, R.  
Direct Submission  
Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarune, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

FEATURES  
source

1..2498  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N"  
/clone="MGC:28503 IMAGE:4188261"  
/tissue\_type="Salivary gland, 10 week old female mouse"  
/clone\_id="NCL CGAP\_SG3"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
55..867  
/codon\_start=1  
/product="similar to region containing hypothetical protein MGCI1352; slingshot"  
/protein\_id="AAH24611.1"  
/db\_xref="GI:19354054"  
/translation="MHYIRISNAEVCWKYKLYFSYIVFWLAGVFLGGLMAWSK  
GVLSDLTKVTRHGIIDPVILWVWFTLGFAGCVGALRENIICLLKFCGAILVLF

CDS

FLLEAVAVIAFLFQDWVRDFREFEFESNIKSYRDIIDIDNLIIDSLKANQCCGAYPE  
DMDLNVFNCSSGASYSREKQVFPSCVDPAPQKVVNTCGDYDRVQLSKSMDERIFT  
KGIOLLEGLEPNRIYIVAGVFALISLDLIFGIFLARILISIDIEAVKAGHNE"

BASE COUNT 542 a 644 c 631 g 681 t

ORIGIN

Alignment Scores:  
Pred. No.: 2,58e-150 Length: 2498  
Score: 1420.00 Matches: 263  
Percent Similarity: 98.52% Conservative: 3  
Best Local Similarity: 97.41% Mismatches: 4  
Query Match: 98.00% Indels: 0  
Gaps: 0

US-09-905-674-2 (1-270) x BC024611 (1-2498)

QY 1 MethIstYrYrArGlyYrSerAsnAlaLysValSerCysTrpYrIlyStYrLeuLeuPhe 20  
|||||  
Db 55 ATGCACATATATAGATACCTGACAGCCGAGGTGAGTGTGATCTTCTTGAGTGGGGCTGTGCGCA 114

QY 21 SerYrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValIleuTrpAla 40  
|||||  
Db 115 AGCTACATATATCTTTTGGCTGGCTGGAGTGTCTTCTTGAGTGGGGCTGTGCGCA 174

QY 41 TrpSerGlyYrGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60  
|||||  
Db 175 TGGAGCGAAAAGGTGTGCTGTCTCGACCTCACCAAGGTGACCCGGTTGCATGGAAATCGAC 234

QY 61 ProValIleuLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80  
|||||  
Db 235 CCCGTCGGCTGGCTGTGATGTGGCGGTGGTGTGATGTTCACACTGGGATTCGCGAGGCTGT 294

QY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValAlaLeu 100  
|||||  
Db 295 GTCCGGGCGCTCCGAGGAACATCTGCTCAAGTTTCTGTGGCGCATTTGCTC 354

QY 101 IlePhePheLeuGlnLeuAlaValAlaValAlaPheLeuPheGlnAspTrpValArg 120  
|||||  
Db 355 ATCTTCTTCTCGGAACGCGCGTGGCTGTGGCTTTTATTTCAGAGCTGGTGAGCA 414

QY 121 AspArgPheArgGlyPhePheGlnSerAsnIleLysSerYrArgAspIleAspLeu 140  
|||||  
Db 415 GACCGGTTCGCGGAATTTCTTCAGAGCAACATCAAGCTATCGGGATTCAGTCACTGCACTG 474

QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaYrGlyProGln 160  
|||||  
Db 475 CAGAACCTCATTTGACTCTCTTCAGAAAGCTAATCAGTGTGCGGGGCTTAAGGCCCTCGAA 534

QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerYrSerArgGlyLysCys 180  
|||||  
Db 535 GACTGGGAGACCTCATGTCTTCACTGCACTGTCAGTGTGCCAGTCAAGCCGACGAAGAAATGT 594

QY 181 GlyValProPheSerCysValProAspProAlaGlnLysValAlaAsnThrGlnCys 200  
|||||  
Db 595 GGGGTACCTTCTCTCTCTGTCGCCAGATCTGCAGAAAAATCCGAAACACACAGTGT 654

QY 201 GlyTyrAspValArgIleGlnIleuLysSerLysTrpAspLeuSerIlePheThrLysGly 220  
|||||  
Db 655 GGCTATGATGTCGGGATTCAGCTGAAAGCAAGGGGATAGTTCACTTTACAAAAGCA 714

QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyAlaPhe 240  
|||||  
Db 715 TGCATCCAGGCTCTGGAAAGCGTGGCTGCCAGAACATCTACATGTGTGCTGTCTTC 774

QY 241 ILeAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260  
|||||  
Db 775 ATTCGCATCTCAGCTGTCAGATTTTGGCATCTCTCGCGAGGACCCCGATCTCAGAC 834

QY 261 ILeGlnAlaValIleAlaGlyHisHisPhe 270  
|||||  
Db 835 ATCGAGGACGATGAAGGACGACATCATCTTC 864

RESULT 6

BC026574 2500 bp mRNA Linear ROD 07-AUG-2002  
 LOCUS BC026574 Mus musculus, clone MGC:36554 IMAGE:4954874, mRNA, complete cds.  
 DEFINITION BC026574  
 ACCESSION BC026574.1 GI:20072352  
 VERSION BC026574.1  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerogamathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2500)  
 REFERENCE Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 CONTACT Contact: MGC help desk  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Jeffrey Green M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www.shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcdpaxll.stanford.edu](mailto:mcdpaxll.stanford.edu)  
 Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plates: 59 Row: a Column: 4  
 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.  
 FEATURES  
 source Location/Qualifiers  
 1..2500  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="FVB/N"  
 /clone="MGC:36554 IMAGE:4954874"  
 /tissue\_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."  
 /clone\_lib="NCI CGAP\_Mam6"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 50..862  
 /codon\_start=1  
 /product="Unknown (protein for MGC:36554)"  
 /protein\_id="AAH26574.1"  
 /db\_xref="GI:20072353"  
 /translation="MHTYRYSNAEYSCMYKYLFSYNTIVFWLQAVFLGVLAMSEK GVLSDLTAVTRLHGIDPVYLVLMGVVMTLGFVCGVALRENICLKFPGCAVLIF FLELAVAVLAFEDPMVDRFREFESNISKSYRDIDQLNIDSLQKANOCCGAYGPE DMDLVNFGSGASYSREKCGPFCSCVDPDAOKVYNOCGYDVRIOLKSKWDEFIFT KGCIOALTEGMLPRNIIYIVAGVFIAISLIQITGIFLARIIDIDIAVYAGHHF"  
 CDS  
 BASE COUNT 545 a 643 c 631 g 681 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,58e-150 Length: 2500  
 Score: 1420.00 Matches: 263  
 Percent Similarity: 98.52% Conservative: 3  
 Best Local Similarity: 97.41% Mismatches: 4  
 Query Match: 98.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-905-674-2 (1-270) x BC026574 (1-2500)  
 Oy 1 MethiStyTyrArgTyrSerAsnAlaLysValSerCysTrpTyrLysTyrLeuAuphe 20  
 Db 50 ATGCACCTATTATATGATACGGAACGCCGAGTCAGCTGCTGATACAGTACTGCTCTTC 109

Oy 21 SerTyrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAla 40  
 Db 110 AGCTACATATATATGCTCTTTGGCTGGCTGGAGTTCTCTTCTTGAGTCGGGCGTGGGCA 169  
 Oy 41 TrpSerGluTyrGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60  
 Db 170 TGGAGCGMAAAGGGGTGTGTGTCTCCGACCTCACCAAGGTGACCCGGTTGCATGTGAATCGAC 229  
 Oy 61 ProValValLeuValLeuMetValGlyValValMetThrLeuGlyPheAlaGlyCys 80  
 Db 230 CCCGTGCTGCTGCTTGTGATGTGGTGGGTGTGTATGTTCACACTGGGATTCGCAAGGCTGT 289  
 Oy 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100  
 Db 290 GTCCGGGCGCCCTCCGAGAGAACATCTGCTGCTCAAGTTTCTGTGGGCCATTGTGCTC 349  
 Oy 101 IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGluAspTrpValArg 120  
 Db 350 ATCTTCTTCTCGAAGCTGGCCGCTGGCGGTGTGGCTTTTATTCMAAGACTGGGTGAGA 409  
 Oy 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeu 140  
 Db 410 GACCGGTTCCGGGAATCTTCGAGAGCAACATCAAGTCTATGGGATGACATCGACTG 469  
 Oy 141 GluAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGlu 160  
 Db 470 CAGAACCTCATATTGCTCCTTCAGAAAAGCATATCATAGTGTGCGGGGCTTCAGGCCCTCGAA 529  
 Oy 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlyLysCys 180  
 Db 530 GACTGGGACCTCATATGCTACTTCACTGCAAGTGTGCCAGCTACACCGAGAAATGT 589  
 Oy 181 GlyValProPheSerCysCysValProAspProIleGlnLysValValAsnThrGlnCys 200  
 Db 590 GGGGTACCTTCTCTCTGCTGTGTCAGATCTGCACAAAAGCTGGAACACACACTGT 649  
 Oy 201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220  
 Db 650 GGCTATGATGTCCTCCGATTCACCTCAAGACGACGATGATGATTCATCTTACAAAAGCA 709  
 Oy 221 CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240  
 Db 710 TGCAATCCAGCGTCTGGAAGGCTGGCTGCCAGCAACATCTACATTTGGCTGTGCTTC 769  
 Oy 241 IleAlaIleSerLeuLeuGlnIlePheGlyTyrIlePheLeuAlaArgThrLysSerAsp 260  
 Db 770 ATTGCCATCTCAGTGTGACAGATTTTGGCATCTTCTGTGGAGACCTGTGATCTCAGAC 829  
 Oy 261 IleGluAlaValLysAlaGlyHisHisPhe 270  
 Db 830 ATCAGGCGAGTGAAGGCGACGATCACTTC 859  
 RESULT 7  
 LOCUS BC002920 2502 bp mRNA Linear PRI 12-JUL-2001  
 DEFINITION BC002920 Homo sapiens, similar to transmembrane 4 superfamily member 9, clone MGC:11352 IMAGE:3954042, mRNA, complete cds.  
 ACCESSION BC002920 GI:12804132  
 VERSION BC002920  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 2502)  
 REFERENCE Submitted (05-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 REMARK



## COMMENT

Contact: MGC help desk  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC/DC/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/nisc.mgc@nih.gov>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
Dierlich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Snyder, B., Stantirlop, S., Thomas, P.J.,  
Tjongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,  
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 15 Row: e Column: 10  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis, similarity but not identity to protein.

FEATURES  
source

1..2502  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:11352 IMAGE:3954042"  
/tissue\_type="skin, melanotic melanoma."  
/clone\_lib="NIR\_MGC\_20"  
/lab\_host="DH10B-R"  
/note="Vector: pORF7"  
63..824  
/codon\_start=1  
/product="similar to transmembrane 4 superfamily member 9"  
/protein\_id="AAH02920.1"  
/db\_xref="GI:12804133."  
/translation="MHYRYRNAKYSWCYKLLFSYNIIFWGVSLDKTVTRMGIDP  
VVLVLMGVVMTLGFAGCVGALRENICILNFCCTIVLIFLELAVAVLAFLEQDMV  
RDREFEFESNKSRYRDLIDNLDLIDSLQAKKCGAGPEMDLNVFNCGSYSR  
KCGVPECCVPDPAKQVYNQCGYDVDRLOKSKDESFITGCIQALDSMLPRNIYI  
VAGVFIAISLDLIEGIFLARTISDIEAVKAGHNP"

BASE COUNT 505 a 711 c 661 g 625 t  
ORIGIN

## Alignment Scores:

Pred. No.: 4,18e-141 Length: 2502  
Score: 1338.50 Matches: 253  
Percent Similarity: 93.70% Conservative: 0  
Best Local Similarity: 93.70% Mismatches: 0  
Query Match: 92.37% Indels: 17  
DB: 9 Gaps: 1

US-09-905-674-2 (1-270) x BC002920 (1-2502)

QY 1 Methistyrtyrargtyrserasnalaalysvalsercystrrpyrlysrtyrleuaphe 20  
|||||  
Db 63 ATGACATATATAGATACCTCTAACGCCAAGCTGAGCTGGTGAACAAGTACCTTTTC 122  
QY 21 SerTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuclYValGlyLeuTPAla 40  
|||||  
Db 123 AGCTACACATCAATCTCTCG----- 143  
QY 41 TrpSerGluysGlyValleuSerAspLeuThrlyValThrArgMetHisGlyIleasp 60  
|||||  
Db 144 -----GGTGGCTGCTCCGACCTCAACAAAGTGACCCGGATGATGAAATGCAC 191  
QY 61 ProValValleuValleuMetValGlyValAlaMetPheThrleuenglyPheAlaGlyCys 80  
|||||  
Db 192 CCTGTGGTGGCTGCTGATGGTGGCTGATGTTCACCGTGGGCTTCGCGGCTGC 251  
QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100

Db 252 GTGGGGCTCTGGCGGAGAAATATCTGCTGTCACTTTTGTGTGGCACATCGTGC 311  
QY 101 IlePhePheLeuGluLeuAlaValAlaValleuAlaPheLeuPheGlnAspTrpValArg 120  
Db 312 ATCTTCTCTCGTAGAGCTGGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 371  
QY 121 AspArgPheArgGluPhePheGlnSerAsnIleCysSerTyrArgAspPheLeu 140  
Db 372 GACCGGTCGCGGAGTCTTTCAGAGCAACATCAAGTCTTACCGGAGCATATGATCTG 431  
QY 141 GlnAsnLeuIleAspSerLeuGlnlyAlaAsnGlnCysCysGlyAlaTyrGlyProGlu 160  
Db 432 CAAACCTCATGACTCTCCCTTCAAAAAGCTTACAGAGTGTGGCGCATATGGCCCTGAA 491  
QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluCys 180  
Db 492 GACTGGAGCTCAACGCTACTTCAATTCAGCGGTCCAGCTACAGCCAGAGAAAGTGC 551  
QY 181 GlyValProPheSerCysCysValProAspProAlaGlnlyValAlaAsnThrGlnCys 200  
Db 552 GGGGTCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611  
QY 201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrlyGly 220  
Db 612 GGATATGATGTCAGATTCACCTGACAGAGCAAGTGGATGATGATCATCTTCACGAAAGGC 671  
QY 221 CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240  
Db 672 TGCAATCCAGCGGCTGGAAGAGCTGCTCCGCGGAACATTTACATTTGGCGGCTGCTTC 731  
QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260  
Db 732 ATCCGCAATCTCGCTGTTCAATATTGGCAATCTTCGCGAAGAGCTGATCTCAGAC 791  
QY 261 IleGluAlaVallyAlaGlyHisHisPhe 270  
Db 792 ATCGAGGCAAGTGAAGCCGCCCATCATCTTC 821  
RESULT 8  
AF065389 1405 bp mRNA linear PRI 28-Apr-2000  
LOCUS  
AF065389 Homo sapiens tetraspan NBT-4 mRNA, complete cds.  
DEFINITION  
AF065389 AF065389.1 GI:3152702  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1405)  
Serru, V., Dessen, P., Bouchaix, C. and Rubinstein, E.  
Sequence and expression of seven new tetraspans  
JOURNAL Biochim. Biophys. Acta 1478 (1), 159-163 (2000)  
MEDLINE 20185353  
PUBMED 10719184  
2 (bases 1 to 1405)  
Rubinstein, E., Serru, V. and Bouchaix, C.  
Direct Submission  
Submitted (14-MAY-1998) INSDERM 0268, 14 av Paul Vallant Couturier,  
Villejuif 94807, France  
FEATURES  
source  
1..1405  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE ID 219547"  
348..1154  
/note="similar to Tspan-5; TM5F"  
/codon\_start=1  
/product="tetraspan NBT-4"  
/protein\_id="AAC17120.1"  
/db\_xref="GI:3152703"  
/translation="MSGKHKGPVSCIKYIFGFNVIFWFLGITFLGIGLMMANKE



Oy		26	PheTpeLeuAagIValValPheLeuGIValGlyLeuTPPaITrPSerGluInysgLy	45
Db		409	TTTTGGTTTGGGAAATACATTCTTGGAATGTGACTGTGGCATGGAAAGAAAAGA	468
Oy		46	ValLeuSerAspLeuThrLysValTrhArgmEtHisGILyleAspProValValLeuVal	65
Db		469	GTTTCGTCCAACATCTCTCCATCACCGATCTCGGGGCTTTGACCAGATTGGCTCTC	528
Oy		66	LeuMetValIGlyValAlValMeCPheThrlLeuGlyPheAlaGlyCySValAlValLeuArg	85
Db		529	CTTGTGGTGGGAGAGTAGTATCTATTCTTTGGATTGTGCAGGCTGCATTGGAGCCCTACGG	588
Oy		86	GLuasnIEcySLeuleuAsnPhenhecySGLyThrlleValleuIIephepheLeuclu	105
Db		589	GAAAAACACTTTCCTCAAGTTTTTCTGTGTTCCGTGGGAATTAATTTCTCTCTGGAG	648
Oy		106	LeuAlaValAlAlaValleuAlAPheLeuPheGlnAsPTPrValArgAspArgPheargGlu	125
Db		649	CTCACTGCCGGAGTTCTAGCATTTGTCTTTCCAAAGACTGGATCAAAAGACCAGCTGATTTTC	708
Oy		126	PhePheGluSerAsnIIelySerTyrrArgSPasPIeaSpLeuGlnAsnLeuIIeasp	145
Db		709	TTTATAACAACAACATCATCAGACATATCGGGATGACATTGATTGCAAAACCTCATAC	768
Oy		146	SerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrglyProGluAsPTPrAspleuAsn	165
Db		769	TTCACCCAGGAATATTGGCAGTGCTGTGGGGCTTTTGGAGCTGATGATGGAACTTAAT	828
Oy		166	ValTyrrPheAsnCysSerGlyAlaSerTyrrSerArgGluLysCySGLYValProPheSer	185
Db		829	ATTTCCTCAATTCATGCACGATTTCCAATCAAGTCCAGAGCGATGTGGGCTTCCATTCTCC	888
Oy		186	CysCySValProAspProAlaGlnLysValAlAsnthrGlnCysGlyTyrrAspValArg	205
Db		889	TGCTCCTACTAAAGATCCGAGAAAGATGTCAACACTCAGTGTGGGTATGATGTCCAGG	948
Oy		206	IleglnLeuLysSerLysTrpAspGluSerIIepheThrlLysGlyCysIIleglnAlaleu	225
Db		949	CAAAACACGAAAGTTGCACAGCAGATTTGTAATCTPCACAGAAAGCGTGTGCCCCAGTTT	1008
Oy		226	GluSerTPrLeuProArgAsnIIelyTrlIleValAlaGlyValPheIIeAlaIIeserLeu	245
Db		1009	GAGAAAGTGTTGGCAGCAGCAATTTAAACATCGTGTGGTATTTCATATGAGCATTCATVG	1068
Oy		246	LeuGlnIIepheGlyIIepheLeuAlaArgThrlLeuIIeserAspIIeglnAlaValLys	265
Db		1069	CTCGAGATATTGGGATATGCTGTGCCCAAGATTGTGTAGCGATATCGAAAGCTGTCAAG	1128
Oy		266	Ala 266	
Db		1129	GCG 1131	
		RESULT 10		
		AF121344		
	LOCUS	AF121344	3175 bp	mRNA linear ROD 01-FEB-2000
	DEFINITION	Mus musculus tetraspanin Tspan-5 (Tspan5) mRNA, complete cds.		
	ACCESSION	AF121344		
	VERSION	AF121344.1	GI:6841032	
	KEYWORDS			
	SOURCE	Mus musculus.		
	ORGANISM	Mus musculus		
		Euarchonta; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;		
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	REFERENCE	1 (bases 1 to 3175)		
	AUTHORS	Garcia-Frigola,C., de Lecea,L. and Soriano,E.		
	TITLE	Mouse Tspan-5 cDNA cloning		
	JOURNAL	Unpublished		
	REFERENCE	2 (bases 1 to 3175)		
	AUTHORS	Garcia-Frigola,C., de Lecea,L. and Soriano,E.		
	TITLE	Direct Submission		
	JOURNAL	Submitted (20-JUN-1999) Dept. of Animal and Plant Cell Biology,		
		University of Barcelona, AV Diagonal 645, Barcelona 08028, Spain		

```
Db 1003 CAGAAACGAGAGTGTACCAAGATTGTATCTACACAAAGGCTGTGTCCCAAGTTT 1062
      ::::: ||::::|
QY 226 GluSerTrpLeuProArgSerIleTyrIleValAlaIleValPheIleAlaIleSerIleu 245
      ||| ||||| ||::: ||::| ||::| ||::| ||::| ||::|
Db 1063 GAGAAAGTGTGCTACAGACATTTTAAACCATTCGTGGTGTATTTTATAGCATTTGCATTG 1122
      ||| ||||| ||::: ||::| ||::| ||::| ||::| ||::|
QY 246 LeuGluIlePheGlyIlePheLeuAlaArgTrpLeuIleSerAspIleGluAlaValys 265
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1123 CTACACATTTTGTGATATGCTCGACACAGAAATTTGTGTAGTGCATTGTAAGCTGTCAAG 1182
      ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 266 Ala 266
      |||
Db 1183 GCT 1185

RESULT 11
LOCUS BC010346 1516 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, RIKEN cDNA 2210021G21 gene, clone MGC:6941
IMAGE:2811935, mRNA, complete cds.
ACCESSION BC010346
VERSION BC010346.1 GI:16307592
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 1516)
REFERENCE 1
  Strausberg, R.
  Direct Submission
  Submitted (05-JUN-2001) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
          DNA Sequencing by: Baylor College of Medicine Human Genome
          Sequencing Center
          Center code: BCM-HGSC
          Web site: http://www.hgsc.bcm.tmc.edu/cdna/
          Contact: amg@bcm.tmc.edu
          Gunaratne, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,
          Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
          Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAX Plates: 5 Row: f Column: 18.

FEATURES
  source
    1..1516
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /map="C57BL/6J"
      /clone="MGC:6941 IMAGE:2811935"
      /tissue_type="Mammary tumor. WAP-1GF alpha model. 7 months
      old, gross tissue."
      /clone_lib="NCI_CGAP_Mam5"
      /lab_host="DH10B"
      /note="Vector: pCMV-SPORT6"
      136..948
      /codon_start=1
      /product="RIKEN cDNA 2210021G21 gene"
      /protein_id="AAH10346.1"
      /db_xref="GI:16307593"
      /db_xref="LocusID:74257"
      /translation="MPGKHQOFDPEVCCGKRYFLGPNIVFWYIALGIGMANG
      EKVLSNLSALTDGGLDPVWLFVYGVGMYSLVPAGCTGLARENTFLKRFSLGL
      IFFELAGITLAFVKDILRDQLNLFNNVKAAYRDDLDLIDFQEWSCCGARG
      PDMWLNITVFNCTDLNPISRCRGVPSFCVDPADVLTNOCGYDIRKLELEOOGSI
```

```
YTRCGVGFPEKMLQDNLIVAGVLGIALIQIFGLCLAONLVSIDIKAVRANW"
BASE COUNT 277 a 417 c 461 g 361 t
ORIGIN
Alignment Scores:
Pred. No.: 1,98e-87 Length: 1516
Score: 861.00 Matches: 151
Percent Similarity: 76.89% Conservative: 52
Best local Similarity: 57.20% Mismatches: 61
Query Match: 59.42% Indels: 0
Gaps: 0
Db: 10

US-09-905-674-2 (1-270) x BC010346 (1-1516)
QY 3 TyrTrpArgTrpSerAsnAlaIleValSerCysTrpTrpTrpLeuPheSerTyr 22
      ::::: ||::| ||| ||||| |||||
Db 148 CACCAAGCATTCACAGAGACCTGAGGTGGCTGGCGGAAATCTCTGTGGCTTC 207
      ||| ||||| ||::: ||::| ||::| ||::| ||::|
QY 23 AsnIleIlePheTrpLeuAlaIleValAlaPheLeuGlyValGlyLeuTrpAlaTrpSer 42
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 208 AACATGTCTCTGTGGTGTGGAGCCCTGTCTGGCCATTGGCCCTGTGGGGGT 267
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 43 GluIysGlyValLeuSerAspLeuThrIysValThrArgMetHisGlyIleAspProVal 62
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 268 GAGAAAGGCGCTTCTCCACATCTCGCGCTGACAGATCGGGCGGCTTGAACCCGTG 327
      ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 ValIleuValIleMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyValGly 82
      ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 TGGCTGTTGTGGTGGGGGAGTGCATGTGCTGTGGCTTGGCGGCTGCATTGGG 387
      ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 AlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValIlePhe 102
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 388 GCCCTCGCGGAAACACCTTCTCTCTCAAGTCTTCTGTCTGTCTCGGCGCTCATCTTC 447
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 103 PheLeuGluIleuAlaValAlaValIleuAlaPheLeuPheGluAsnTrpValArgAspArg 122
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 448 TTCTGTGAGCTGGCGGGGGGATCTGCTGTCTGTCAAGGATTTGATCCGAGACCA 507
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 123 PheArgGluPhePheGluSerAsnIleIleSerTrpTrpArgAspIleAspLeuGluAsn 142
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 508 CTATATCTCTTCATCAACACAAATGTCAAGACCTACCGGATGTATCTCGACTTCGAAC 567
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 LeuIleAspSerLeuGluIleuAlaAsnGlnCysGlyAlaTrpGlyProGluAspTrp 162
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 568 CTATTCGACTTGTGCTCAGAAATAGTGTCTGTCTGTGTGTCCGAGGCCCATAGTGTG 627
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 163 AspLeuAsnValIleTrpAsnGlySerGlyAlaSerTrpSerArgGluIysCysGlyVal 182
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 628 AACCTCAATATCTTCAACTGCATGCATGTGAACCAAGCCGTGAGCCGTGTGGGCTG 687
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 183 ProPheSerCysValProAspProAlaGlnIleValAlaAsnThrGlnCysGlyTrp 202
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 688 CCTTTTCTCTCTGTGTAGGACCCAGCGGAAATATCTCTCAATATCCCATGTGTGCTAT 747
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 203 AspValArgIleGlnIleuIysSerIysTrpAspGluSerIlePheThrIysGlyCysIle 222
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 GACATTCGACTCAACATGAGACTCGACAGAGGCTCCATCTACACCAAGGCTGTGCTG 807
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 223 GlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaIleValAlaIle 242
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 GGCCAGTTTGAAGATGCTACAGCAACACTGATCTGTGGCCGGGCTTGTGGGG 867
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 243 IleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlu 262
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 868 ATGCGCTCTCTCGAGATATTTGGCTCTGTGCTGCGCCAGAACTCGTAGTGCATCAAG 927
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 263 AlaValIysAla 266
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 928 GCGGTGAAGGCC 939

RESULT 12
AX420468 813 bp DNA linear PAT 18-JUN-2002
LOCUS
```

```

DEFINITION      Sequence 3 from Patent WO0216603.
ACCESSION       AX420468
VERSION         AX420468.1  GI:21524616
KEYWORDS
SOURCE          human.
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS         Leiby, K.R.
TITLE           23228, a human tetraspanin family member and uses thereof
JOURNAL         Patent: WO 0216603-A 3 28-FEB-2002;
                Millennium Pharmaceuticals, Inc. (US)
FEATURES
source          1..813
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT      151 a 233 c 235 g 194 t
ORIGIN
Alignment Scores:
Pred. No.:      1 94e-87      length:      813
Score:          858.00      Matches:      154
Percent Similarity: 75.76%      Conservative: 46
Best Local Similarity: 58.33%      Mismatches: 64
Query Match:    59.21%      Indels:      0
DB:             Gaps:      0

US-09-905-674-2 (1-270) x AX420468 (1-813)
QY      3 TyrTyrAgtYrSerAsnAlaLysValSerCysTrrPtyrLysTyrLeuLeuPheSerTyr 22
        :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      13 CACCACACATTTCCAGAACCTGAGTGGCTGCTGGCGGAAATACCTCTCTGGCTTC 72

QY      23 AsnIleIlePhePheLeuAglValValPheLeuGlyValGlyLeuTrrPalatrrSer 42
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      73 AACATTGCTCTTGGTGGTGGAGCCCTGTTCTCGCTGCTGCTGCTGCTGCTGCTGCTG 132

QY      43 GluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProVal 62
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      133 GAAAGGGCTTCTTCGAACATCTCAGCGCTGACAGATCTGGAGGCCCTTGACCCCGTG 192

QY      63 ValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAglCysValGly 82
        ||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      193 TGGCTGTTGGTAGTAGTAGGAGCGCTCAGTGGCTGCTGCTGCTGCTGCTGCTGCTG 252

QY      83 AlaLeuArgLysAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhe 102
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      253 GCCCTCCGGAGAACCTTCCTCGCTCAAGTTTCTCCGTCTCTCGCTCAATCTTC 312

QY      103 PheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGluAspTrrPalatrrSpar 122
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      313 TTCTCTGAGCTGGCAACAGGATCCCTGCTTGTCTTCAAGAGCTGGATTCGAGACCA 372

QY      123 PheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGlnAsn 142
        ||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      373 CTCAACCTCTTCATCAACAACACGTCAGGCTACCGGAGACGACATTGACCTCCAGAAC 432

QY      143 IleuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaIleArgGlyProGluAspTrr 162
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      433 CTGATTGACTTGGCTGAGAAATACCTGCTCTGTGGAGCCCGAGGCCCAATACCTGG 492

QY      163 AspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCysGlyAla 182
        :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      493 AACCTCAATATCTACTTCAACTGACATGATTCGAACCCGAGCGGAGGCTGCGGGGTG 552

QY      183 ProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCysGlyTyr 202
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      553 CCTTCTCTGCTGCTCAGGAGCCCTGGGAGGATGCTCTCAACACCAAGTGGCTAC 612

QY      203 AspValArgIleGlnLeuLysSerLysTrrPaspGluSerIlePheThrLysGlySyle 222
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      613 GAGCTCGGCTCAAACTGGAGCTGGAGACGACGAGGCTTCATCCACCAAGGCTGCGTG 672

```

```

QY      223 GlnAlaLeuGluSerTrrPheuProArgAsnIleTrrIleValAlaGlyValPheIleAla 242
        ||| ||||| ||::: :::: :::: :::: :::: :::: ::::
Db      673 GGCACGTTTGAGAAAGTGTGCGAGCAACCTGATTTGTGTGGCGGAGCTTCATGGCC 732

QY      243 IleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgTrrLeuIleSerAspIleGlu 262
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      733 ATCGCCCTTCCTCAATCTTTGGCATTCCTGCTGGCCGAGAACCTGTGATGACATCAAG 792

QY      263 AlaValLysAla 266
        ||||| ::::
Db      793 GCAGTGAAGGCC 804

RESULT 13
AX420466
LOCUS          AX420466 3184 bp DNA linear PAT 18-JUN-2002
DEFINITION    Sequence 1 from Patent WO0216603.
ACCESSION     AX420466
VERSION       AX420466.1  GI:21524614
KEYWORDS
SOURCE        human.
ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS       Leiby, K.R.
TITLE         23228, a human tetraspanin family member and uses thereof
JOURNAL       Patent: WO 0216603-A 1 28-FEB-2002;
                Millennium Pharmaceuticals, Inc. (US)
FEATURES
source        1..3184
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD35489.1"
                /db_xref="GI:21524615"
                /translation="MGKKHOFDEPEVCGCGKFLFEGNIVFWLCAFLAIGLWAG
                EKGVLNLSIALTDIGLDIPWLFVVGVSIVGFGCGALNENFLKFESEVLAGL
                IFLELATGTLAEVFKXMDIRDQNLNLFNNNVKVAIRDIDQNLIDFAQEWSCGARG
                PNDMLNIVFNCTDLNPSRERGVPSCCVBPDAEDVLANTCGYDRLKLEIQGFEI
                HTKGCVGQEFKMLQDNLIYVAGVPMGIALQIIGICLAQNLVSIDIAVNAANM"

BASE COUNT    675 a 883 c 945 g 681 t
ORIGIN
Alignment Scores:
Pred. No.:      1.13e-86      length:      3184
Score:          858.00      Matches:      154
Percent Similarity: 75.76%      Conservative: 46
Best Local Similarity: 58.33%      Mismatches: 64
Query Match:    59.21%      Indels:      0
DB:             Gaps:      0

US-09-905-674-2 (1-270) x AX420466 (1-3184)
QY      3 TyrTyrAgtYrSerAsnAlaLysValSerCysTrrPtyrLysTyrLeuLeuPheSerTyr 22
        :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      180 CACCACATTTCCAGAACCTGAGTGGCTGCTGGCGGAAATACCTCTCTGGCTTC 239

QY      23 AsnIleIlePhePhePLeuAglValValPheLeuGlyValGlyLeuTrrPalatrrSer 42
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      240 AACATTGCTCTTGGGCTGCTGGAGCCCTGCTCTGCTGCTGCTGCTGCTGCTGCTG 299

QY      43 GluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProVal 62
        ||||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      300 GAGAAGGCGTTCCTCGAACATCTCAGCGCTCAGACATCTGGAGAGCCCTTGACCCCGTG 359

QY      63 ValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAglCysValGly 82
        ||| :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      360 TGGCTGTTGGTAGTAGTAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 419

QY      83 AlaLeuArgLysAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhe 102

```









GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 19, 2002, 06:14:57 : Search time 2003 Seconds

(without alignments)  
2183.116 Million cell updates/sec

Title: US-09-905-674-2

Perfect score: 1449

Sequence: 1 MHYRYSNAKVCWYKILF.....IFLARTLISDEAVKAGHHF 270

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO/spool/US0905674/runat\_16122002.111846\_6769/app\_query.fasta.1.455  
-DB=EST -QFMT=fastap -SUPEX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCCALIG=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MTLEN=0 -MAXLEN=2000000000  
-USER=US0905674.ecgn.1.1.899\_@runat\_16122002.111846\_6769 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOC -DEV=TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1436	99.1	899	14	BQ212057
2	1430	98.7	1104	14	BM923308
3	1383	95.4	866	12	BG674208
4	1324	91.4	1125	14	BQ067138
5	1304	90.0	883	12	BG753541
6	1287.5	88.9	857	13	B1835459
7	1286	88.8	836	13	B1905852
8	1276	88.1	831	13	B1663945
9	1275	88.0	742	14	BQ572112
10	1271	87.7	951	9	AL519790
11	1246.5	86.0	797	13	BG913227
12	1246.5	86.0	870	14	BQ948800
13	1244	85.9	755	12	BG765719
14	1228	84.7	810	12	BG760633
15	1223	84.4	770	13	BG974843
16	1218	84.1	940	14	BQ943303
17	1197.5	82.6	845	12	BG862723
18	1197.5	82.6	845	13	B1853594
19	1188	82.0	764	13	B1853594
20	1181.5	81.5	1013	13	BG915134
21	1173	81.0	724	12	BG770453
22	1171	80.8	730	13	B151673
23	1165	80.4	728	13	B1556034
24	1156	79.8	824	13	B1912573
25	1144	79.0	925	9	AL539537
26	1133	78.2	970	12	BG760833
27	1128	77.8	694	13	BM051048
28	1122	77.4	751	12	BG246551
29	1120	77.3	841	12	BF026182
30	1118.5	77.2	924	12	BE895131
31	1090	75.2	924	13	B1251280
32	1088	75.1	1001	12	BF027356
33	1083	74.7	975	12	BF785494
34	1081	74.6	973	12	BG479961
35	1070	73.8	794	12	BF537170
36	1055.5	72.8	981	12	BG176342
37	1050.5	72.5	855	12	BF099030
38	1047	72.3	744	10	BE276487
39	1045	72.1	710	10	BE276329
40	1042	71.9	743	10	BE410894
41	1034.5	71.4	897	12	BF785457
42	1031.5	71.2	671	10	BE388047
43	1027	70.9	780	13	B1463709
44	1022.5	70.6	809	10	BE384608
45	1017	70.2	635	13	BG974146
	1011.5	69.8	597	10	BE408222

#### ALIGNMENTS

RESULT 1  
LOCUS BQ212057  
DEFINITION BQ212057 899 bp mRNA linear EST 02-MAY-2002  
ACCESSION BQ212057  
VERSION BQ212057.1 GI:20391921  
KEYWORDS  
SOURCE human.  
ORANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL COMMENT

Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: ATCC/DCMD/DFP  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLAM13331 row: a column: 01  
 High quality sequence stop: 675.

## FEATURES

source

Location/Qualifiers

1. .899

/organism="Homo sapiens"

/db.xref="taxon:9606"

/clone="IMAGE:6061320"

/clone.lib="NIH\_MGC\_72"

/tissue\_type="melanotic melanoma"

/lab.host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: PCMV-SPORT6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT

190 a

234 c

259 g

214 t

2 others

ORIGIN

## Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

14

Gaps:

0

Matches:

267

Mismatches:

2

Indels:

0

US-09-905-674-2 (1-270) x BQ212057 (1-899)

QY 1 MethisTyrTyrArgTyrSerAsnAlaLysValSerCysTrpTyrLysTyrLeuLeuPhe 20  
 |||||

Db 5 ATGCACTATTATGATGACTCTAAAGCCAGGTCAGCTGCTGATACAGTCACTCTTTTC 64  
 |||||

QY 21 SetTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40  
 |||||

Db 65 AGCTACAAATCATCTTCTGGTGGCTGGATGTTGTTCTTCTTGAGTGGGCTGTGGCA 124  
 |||||

QY 41 TrpSerGlyLysGlyValLeuSerAspLeuThrLysValThrArgMethIleAsp 60  
 |||||

Db 125 TGGAGGAAAAGGCTGCTGTCCGACCTCACCAAGTACCCGATGATGGAATCGAC 184  
 |||||

QY 61 ProValIleValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80  
 |||||

Db 185 CCTGTGCTGCTGCTGATGTGGGCTGGTGGATGTTCACTCGGGTGGCTGCTGC 244  
 |||||

QY 81 ValGlyAlaLeuArgLysAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100  
 |||||

Db 245 GTGGGGGCTCTGGGAGATATGCTGCTGCTCACTTTTCTTGCGACCATCTGCTC 304  
 |||||

QY 101 IlePhePheLeuGlyLeuValAlaValAlaPheLeuPheGlnAspTrpValArg 120  
 |||||

Db 305 ATCTTTCTCTGAGCTGCTGTGGCCGCTGCTGCTTCTTCCAGGACTGGGTGAGG 364  
 |||||

QY 121 AspArgPheArgGlyPhePheGlnSerAsnIleLysSerTyrArgAspAspIleAspLeu 140  
 |||||

Db 365 GACCGGTTCCGGAGATTCTTCAGAGCAACATCAAGTCTACCGGAGCATGATCATCG 424  
 |||||

QY 141 GlnAsnIleLeuAspSerLeuGlnLysAlaAsnGlnCysGlyAlaIleArgProGlu 160  
 |||||

Db 425 CAAACCTATGACTCCCTTCAAGAAAGCTAAACAGTGTGGGCGATATGGCCCTTAA 484  
 |||||

QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlyLysCys 180  
 |||||

Db 485 GACTGGACCTCAACGCTTACTTCAATTGCGAGGCTGCCAGTACGCCGAGAGAGATGC 544  
 |||||

QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCys 200  
 |||||

Db 545 GGGGTCCCTTCTCCTGCTGGCTGCCAGATCTTCGCGCAAAAGTTGGAAACACACACTGT 604  
 |||||

QY 201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGlnSerIlePheThrLysGly 220  
 |||||

Db 605 GGATATGATGTGCAGATTCACACTCAAGAGCAAGGAGTGTGCTCATCTTCACCAAAAGGC 664  
 |||||

QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240  
 |||||

Db 665 TGCATCCAGGGGCTGGAAAGCTGGCTCCCGGGAACATTATTCATTTGCTGGCTTTTC 724  
 |||||

QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260  
 |||||

Db 725 ATGCGCATCTCGCTGTTCAGATATTTGGCATCTTCTCGGCAAAAGCGTATCTCAGAC 784  
 |||||

QY 261 IleGlnAlaValLysAlaGlyHisHisPhe 270  
 |||||

Db 785 ATCNAAGCAGTGAAGGCGCCATCCTTC 814  
 |||||

RESULT 2

LOCUS BM923308 1104 bp mRNA linear EST 12-MAR-2002

DEFINITION ACENCCURT\_6625806 NIH\_MGC\_116 Homo sapiens CDNA clone IMAGE:5758901

ACCESSION BM923308

VERSION BM923308.1 GI:19373687

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1104)

NIH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM12803 row: h column: 06

High quality sequence start: 7

High quality sequence stop: 700.

Location/Qualifiers

1. .1104

/organism="Homo sapiens"

/db.xref="taxon:9606"

/clone="IMAGE:5758901"

/clone.lib="NIH\_MGC\_116"

/lab.host="DH10B"

/note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH-MGC Library."

BASE COUNT 218 a 315 c 319 g 251 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1 76e-164 Length: 1104

Score: 1430.00 Matches: 269

Percent Similarity: 99.26% Conservative: 0

Best Local Similarity: 99.26% Mismatches: 1  
 Query Match: 98.69% Indels: 1  
 DB: 14 Gaps: 0

US-09-905-674-2 (1-270) x BM923308 (1-1104)

QY 1 Methistyrtyrarglytyrserasnalaalysvalsercysrtpttyrlystyreuleuphe 20  
 |||||||  
 DB 72 ATGCACATATATAGATACCTAAGCCCAAGGTCACCTGCTGTACAGTACCTCTTTTC 131  
 QY 21 Sertyrasnillelephetrpleunlaaglyvalalpheleuglyvalglyeunrpla 40  
 |||||||  
 DB 132 ACCTACACATCATCTTCTGTGGTGGAGTGTCTTCTTGGAGTGGGCTGTGGGA 191  
 QY 41 TTPserglytyrvalleyaleuseraspleunrlyrsvalthrarmethisglylleasp 60  
 |||||||  
 DB 192 TGGAGCAAAAGAGTGTGTGTCCGACCTCACCAAGATGACCGGATGATGAAATCGAC 251  
 QY 61 Provalvalleuvalleuvalleyalvalmetpetherleuglyphealaaglycys 80  
 |||||||  
 DB 252 CCTGT 311  
 QY 81 Valglyalaleuarglylunsnilecysleuleuasnphepcysglythrillevalleu 100  
 |||||||  
 DB 312 GTGGGGGCTGTGGGGAATATCTGTGCTCACTTTTCTGTGGACATCGTGTCTC 371  
 QY 101 llephphelenuleuvalaalaalaleuvalleuvalpheleuphegnlasptypvalar 120  
 |||||||  
 DB 372 ATCTTCTCTGTGAGT 431  
 QY 121 Aspariphearglyluphepegusasnillelsercysrttyrargaspasplleaspleu 140  
 |||||||  
 DB 432 GACCGGTTCGGGAGTCTTCTGTGAGCAACATCACTACCGGACATATCGATCTG 491  
 QY 141 Gluasleuulleaspserrleuglylunslalaasnglncyscysglvalatryglyproglu 160  
 |||||||  
 DB 492 CAHAACCTCATGACCTCCTTCAGAAAGCTAACCAAGTGTGTGTGTGTGTGTGTGTGA 551  
 QY 161 Asptirpaspheuasnyaltrypheasnrcysserglyalasercysrttyrarglylucys 180  
 |||||||  
 DB 552 GACTGGGACCTCAACGCTACTTCAATTTGACGGGTGCCAGCTACAGCCGAGAGAGTGC 611  
 QY 181 Glyvalprophesercyscysvalproaspproalaglnlyvalalasnthrincys 200  
 |||||||  
 DB 612 GGGGTCCCTTCTCTCTGT 671  
 QY 201 Glytyrtraspvalarglylleuglyserlystraspeluserlelephetrlysgly 220  
 |||||||  
 DB 672 GGAATATATGTCAAGATTCAGCTGAGAGCAAGTGGATGATCCATCTTCACGANAAGC 731  
 QY 221 Cyslleuglinalaenguuserttrypleuoproarasnillefyrillevalaaglyvalphe 240  
 |||||||  
 DB 732 TGCATTCAGGGCGGTGAAGACCTGGCTCCCGGAACTTTACATTTGGCTGGCGTCTTC 791  
 QY 241 llealalleserleuenglnillepheglyllepheleuualarqthrleuilleserasp 260  
 |||||||  
 DB 792 ATCCCAATCTCGCTGTGTGAGATATTTGGCATCTTCTGTGCAAGGACGCTGATCTCAGAC 851  
 QY 261 lleglu-AlaValalysalaglyhshlspe 270  
 |||||||  
 DB 852 ATCGAAGGCACTGAAGCCGGCCATCTCTTC 882

RESULT 3  
 BG674208 866 bp mRNA linear EST 01-MAY-2001  
 LOCUS 602619868F1 NCL\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4745512 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG674208  
 VERSION BG674208.1 GI:13905604  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 866)  
 AUTHORS NIH-MGC  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: James Cleaver, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA  
 Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LNL0592 row: 9 column: 17  
 High quality sequence stop: 821.  
 Location/Qualifiers  
 1..866  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4745512"  
 /clone\_lib="NCL\_CGAP\_Skn3"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-SPORT6; site\_1: NotI;  
 site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.5kb. Library constructed by Life  
 Technologies. Note: this is a NCL\_CGAP Library."  
 BASE COUNT 176 a 229 c 255 g 206 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 7.02e-159 Length: 866  
 Score: 1383.00 Matches: 262  
 Percent Similarity: 98.87% Conserved: 0  
 Best Local Similarity: 98.87% Mismatches: 2  
 Query Match: 95.45% Indels: 1  
 DB: 12 Gaps: 0

US-09-905-674-2 (1-270) x BG674208 (1-866)

QY 7 Serasnalaalysvalsercysrttyrlystyreuleupheserttyrasnillelephe 26  
 |||||||  
 DB 2 TCTAACGCCAAGTCAAGT 61  
 QY 27 TTPleualaglyvalalpheleuglyvalglyleuTPalatrpsertglyglyval 46  
 |||||||  
 DB 62 TGGT 121  
 QY 47 Leuseraspleunrlyrsvalthrarmethisglylleaspprovalvalleuvalleu 66  
 |||||||  
 DB 122 GTGTCCGACCTCACCAAGATGACCCGGATGATGGAATCGACCCCTGTGTGTGTGTGT 181  
 QY 67 Metvalglyvalalmetpetherleuglyphealaaglycysvalglyalaleuargly 86  
 |||||||  
 DB 182 ATGTGTGGGCGGT 241  
 QY 87 Asnillecysleuleuasnphepcysglythrillevalleuullepheleuglyleu 106  
 |||||||  
 DB 242 AATATCTGTCTCAACTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 301  
 QY 107 AlaValalValleuvalpheleuphegnlasptypvalaraspasplleaspleu 126  
 |||||||  
 DB 302 GCTGTGGCGGT 361  
 QY 127 PhegluserasnillelsercysrttyrargaspasplleaspleuGlnsnleuillespser 146  
 |||||||  
 DB 362 TTGAGAGCAACATCAAGCTTACCGGACGATATCGATCTCAAAACCTCATTCACATCC 421  
 QY 147 Leuglnlyalalasnnglncyscysglvalatryglyprogluaspaspleuasnyal 166  
 |||||||  
 DB 422 CTTCAGAAAGCTAACCAAGT 481  
 QY 167 TyrPheasnrcysserglylaserttyrserarglylucyscysglyvalprophesercys 186

```

|||||
Db 482 TACTCAATTGCAGCGGTGCACCTACAGCCGAGAGAGTGGGGGTCCCTTCCTCC 541
|||
QY 187 CysValProAspProAlaGlnValValAsnThrGlnCysGlyTyrAspValArgIle 206
|||
Db 542 TGGCTGCGAGATCTCTCGCAAAAAGTTGTAACACACATGTGGATATATGTCAGGATT 601
|||
QY 207 GlnLeuLysSerIleTyrAspGluSerIlePheThrLysGlyCysIleGlnAlaLeuGlu 226
|||
Db 602 CACCTGACAGGCAAGTGGGATGATGCTCATCTTCACGAAAGGCTGCATCCAGCGCTGGAA 661
|||
QY 227 SerTyrLeuProArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeu 246
|||
Db 662 AGCTGGCTCCCGGCAACATTTACATTTGCTGCTGGCTCTTCATCGCCATCTCGCTGTTG 721
|||
QY 247 GlnIlePheGlyIlePheLeuAlaArg-ThrLeuIleSerAspIleGlnAlaValLysAl 266
|||
Db 722 CAGATTTTGGCATCTCTTCGCAAGGAGCGCTGATCTCAGCATCGAGCGAGTGAAGGC 781
|||
QY 266 aclyhshshspshe 270
|||
Db 782 CGAGCATCACTTC 794
|||
RESULT 4
BQ067138 1125 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT_6767382 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5751493
DEFINITION BQ067138 5', mRNA sequence.
ACCESSION BQ067138
VERSION BQ067138.1 GI:19896184
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1125)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRAM12784 row: C column: 14
High quality sequence stop: 587.
Location/Qualifiers
1. 1125
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5751493"
/clone_1ib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 239 a 333 c 297 g 256 t
ORIGIN
Alignment Scores: 1.86e-151 Length: 1125
Pred. No.: 1324.00 Matches: 260
Score: 5

```

```

Percent Similarity: 95.658 Conservative: 4
Best Local Similarity: 94.208 Mismatches: 6
Query Match: 91.37% Indels: 6
DB: 14 Gaps: 0
US-09-905-674-2 (1-270) x BQ067138 (1-1125)
QY 1 MethsIstYrYrAqYrYrSerAsnAlaLysValSerCysTyrIleTyrLeuLeu 20
|||
Db 93 ATGACATATTATAGTACTCTTAACGCCAAGTCAAGTCAAGTGTGTGACAGTCACTCCCTTC 152
|||
QY 21 SerTyrAsnIleIlePheThrLeuAlaGlyValValPheLeuGlyValIleTyrPala 40
|||
Db 153 AGCTACACATCATCTTCTGTGTGGTGGAGTGTCTTCTTGTGAGTGGGGCTGTGGCA 212
|||
QY 41 TrpSerGlnIleGlyValLeuSerAspLeuThrLysValThrArgMetIshGlyLeasp 60
|||
Db 213 TGGAGCGAAAGGGGTGTGTGTGCGACCTCACCAGATGACCCGAGTGCATGGATTCGAC 272
|||
QY 61 ProValValLeuValIleuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys 80
|||
Db 273 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 332
|||
QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
|||
Db 333 GTGGGGCTCTCGCGGAAATATCTGCTTCTCACTTTCTGTGTGACCATGTGTCTC 392
|||
QY 101 IlePhePheLeuGlnLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTyrValArg 120
|||
Db 393 ATCTTCTTCTCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 452
|||
QY 121 AspArgPheArgGluPhePheGlnSerAsnIleLysSerTyrArgAspIleLeu 140
|||
Db 453 GACGGGTTCGGGAGTCTTCGAGACCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 512
|||
QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlyCysGlyAlaTyrGlyProGlu 160
|||
Db 513 CAAACCTCATGCAGTCCCTTCAGAAAGCAACCACTGCTGTGTGTGTGTGTGTGTGTGTGT 572
|||
QY 161 AspTyrAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180
|||
Db 573 GACTGGGACCTCAACCTCACTTCAATTCACGGGTGCCAGCTACAGCGAGAGAGAGTGC 632
|||
QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
|||
Db 633 GGGGTCCCTTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 692
|||
QY 201 GlyTyrAspValArgIleGlnLeuLysSerIleTyrAspGluSerIlePheThrIleGly 220
|||
Db 693 GGTATGATGTCTCAAGTATCACTGAGAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 752
|||
QY 221 CysIle-GlnAlaLeuGlnSerTyrLeuProArg-AsnIleTyrIleVal-AlaGlyVal 239
|||
Db 753 TGTATTCACGGGCTGTGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 812
|||
QY 240 PheIleAlaIleSerLeu-LeuGlnIlePheGlyIlePhe-LeuAlaArgThrIleLys 259
|||
Db 813 TTCAATGCCATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 872
|||
QY 259 erAspIleGlnAla-ValLysAlaGlyhshspshe 270
|||
Db 873 CAACATTCACGAGCATTAAGCGCGGCACTTCATTC 908
|||
RESULT 5
BQ753541 883 bp mRNA linear EST 15-MAY-2001
LOCUS 602733110P1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4876370 5',
DEFINITION mRNA sequence.
ACCESSION BQ753541
VERSION BQ753541.1 GI:14064194
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 883)  
 NIH-MGC <http://mhc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LCM1759 row: d column: 03  
 High quality sequence stop: 786.

FEATURES  
 source  
 1..883  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4876370"  
 /clone\_lib="NIH\_MGC\_43"  
 /tissue\_type="normal pigmented retinal epithelium"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pOT7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Library constructed by Ling Hong in the laboratory of Gerald W. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

BASE COUNT 174 a 238 c 264 g 207 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,61e-149 Length: 883  
 Score: 1304.00 Matches: 262  
 Percent Similarity: 95.27% Conservative: 0  
 Best Local Similarity: 95.27% Mismatches: 8  
 Query Match: 89.99% Indels: 7  
 DB: 12 Gaps: 0

US-09-905-674-2 (1-270) x BG753541 (1-883)

OY 1 MethISTYRTYRGTYSerASnAlaYSValSerCysTTPYRlySlyLeuLeuPhe 20  
 DB 51 ATGCACATATATAGTACTCTACGCCAAGGTCAGCTGCTGTACAACTCTCTTTTC 110  
 OY 21 SerTYRAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40  
 DB 111 AGCTACACATCATCTCTCGTTGGCTGGAGTGTCTCTCTTGAGAGTGGGCTGGGCA 170  
 OY 41 TTPSerGlyulysGlyValLeuSerAspIleThrIysValIThrArgMethISglyIleasp 60  
 DB 171 TGGACCGAAAGGCTGCTGTCGACCTCACCAAGACCCGATCATGAAATGCAC 230  
 OY 61 ProValIleValLeuMetValGlyValIleMetPheThrIleGlyPheAlaGlylys 80  
 DB 231 CCTGTGGTGGCTGCTCTATGATGCTGGGCTGTATGATGCTGGGCTTGGCGGCTGC 290  
 OY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100  
 DB 291 GTGGGGGCTTGGCGGAATATCTGCTCAACTTTTCTGTGGCACCATGCTGCTC 350  
 OY 101 IlePhePheLeuGlyLeuAlaValAlaValAlaPheLeuPheGlnAspTrpValArg 120  
 DB 351 ATCTTCTTCTGGAGCTGGTGGCTGGCTGGCTTCCTTCCTTCAGAGCTGGGTGAGG 410  
 OY 121 AsparArgPheArgGlyPhePheGlySerAsnIleIysSerTYRArgAspAspIleAspLeu 140  
 DB 411 GACCGGTTCCGGAGATTCTTCGAGAGCAACATCAAGTCTTACCGGAGCATATGATCTG 470

OY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrglyProGlu 160  
 DB 471 CAAAACCTCATGCACTCCCTTCAGAAAGCTAACAGTCTGTGGCGCATATGGCCCTGAA 530  
 OY 161 AspTrpAspLeu-AsnValIlePheAsnCysSerGlyAlaSerIysSerArgGlyulysCys 180  
 DB 531 GACTGGAGCTCAACGCTACTTCATTAATTCAGCGGTGCTACAGCCGAGAGAGTG 590  
 OY 180 sGlyValProPheSerCysValPro-AspProAlaGlnLysValAlaAsnThrGln 200  
 DB 591 CGGGTCCCTCTCTCTCTCTCTGCGCCAGATCTCTGCGCAAAAGT-CTGACACACAGT 649  
 OY 200 ySGlyTYRAspValArgIleGlnLeuIysSerlySTPAspGlySerIlePheThrIys- 219  
 DB 650 GTGGATGAGATGATCAGGATTCAGCTGAAGACCAAGTGGAGAGTCCATCTTCAGCAACA 709  
 OY 220 GlyCysIleGlnAlaLeuGlu-SerTrp-LeuProArgAsnIleTyrlleValAlaGly 239  
 DB 710 GGCTGCATCCAGCGCTGGAAAAGCTGCCCTCCCGGACATTACATTGTGGCTGGCG 769  
 OY 239 alPheIleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIles 259  
 DB 770 TCTTCATCGCATCTCGCTGTGGCCGATATGGCATCTTCTGCGAGCGGTATCT 829  
 OY 259 eraSpIleGlnAlaValIysAlaGlyHisHisPhe 270  
 DB 830 CAGACATCGAGGAGTGAAAGCCGG--CATCACTTC 862

RESULT 6 857 bp mRNA linear EST 04-OCT-2001  
 B1835459  
 LOCUS 603087572F1 NIH\_MGC\_120 Homo sapiens CDNA clone IMAGE:5226856 5',  
 DEFINITION  
 mRNA sequence.  
 ACCESSION B1835459  
 VERSION B1835459.1 GI:15947009  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 857)  
 NIH-MGC <http://mhc.nci.nih.gov/>.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLAM1570 row: o column: 17  
 High quality sequence stop: 840.

FEATURES  
 source  
 1..857  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5226856"  
 /clone\_lib="NIH\_MGC\_120"  
 /lab\_host="DH10B"  
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb. Insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC library."

BASE COUNT 159 a 236 c 246 g 216 t



|||||  
Db 351 ATCTTCTTCCTGAGCTGCTGCTGCGCGCTGCGCTTCTGCTTCCAGAGAGCTGGTAGG 410  
OY 121 ASPATGPhaerGluPhepGluSerAsnIleYsserTYrArAspAspIleAspIleu 140  
Db 411 GACCGGTTCGGAGTCTTCCAGAGACATCAAGTCTTACCAGCATATTCGATCTG 470  
OY 141 GlnAsnIleAspSerIleuGlnIlysaAsnGlnCysGlyAlaTyrGlyProGlu 160  
Db 471 CAAACCTCATGACCTCCCTTCCAGAAAGTACCAAGTGTGTGGGCGCATATGCGCTTAA 530  
OY 161 ASPPTPASPleuAsnValTYrPheAsnCYsserGlyAlaSerTYrSerArGluIlyscys 180  
Db 531 GACTGGACCTCAAGCTACTTCAATTGCACGCGTGCACGTACAGCGGAGAAAGTGC 590  
OY 181 GlyValProPheSerCysValProAspProAlaGlnIlyValAlaAsnThrGlnCys 200  
Db 591 GGGGTCCCTTCTCTGCTGCGCGCATGCTGCGCAAAAGTGTGAACACACAGTGT 650  
OY 201 GLYTYrAspValArgIleGlnIleuIlysserIlystrP-AspIuserIlePheThrIlysgl 220  
Db 651 GGATATGATGTCAGATTCAGCTGAAGACAGTGGGATGATGTCATCTTCACGAAAG 710  
OY 220 yCys-IleGln-AlaIleuGlnSerTrpleu-ProArGAsnIleTYrIleValAlaGlyVa 239  
Db 711 CTGCATCCAAAGCGCTGGAAGCTGCTGCCGCGGAACATTTACATTTGCTGCGCGT 770  
OY 239 IPhelIleAlIleSerIleuGlnIle-PheGlyIlePheIleuAlaArgThrIleuIles 259  
Db 771 CTTTCATCCGATCTCGCGTGTGCAATATTTGGGCAATCTCTCCCTGGAAGGCGCTGATCT 830  
OY 259 er 259  
Db 831 CC 832  
RESULT 8  
B1663945 831 bp mRNA linear EST 12-SEP-2001  
LOCUS 60328609F1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:5322531 5',  
DEFINITION mRNA sequence.  
ACCESSION B1663945  
VERSION B1663945.1 GI:15578178  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS 1 (bases 1 to 831)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-rc@mail.nih.gov  
Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: LLM11818 row: b column: 04  
High quality sequence stop: 789.  
location/Qualifiers  
1..831  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5322531"  
/clone\_lib="NCI\_CGAP\_Mam6"  
/sex="female, virgin"  
/tissue\_type="infiltrating ductal carcinoma"  
/dev\_stage="5 months"  
/lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-Sport6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Jeffrey Green, M.D., NIH"  
BASE COUNT 160 a 218 c 239 g 214 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,97e-146 Length: 831  
Score: 1276.00 Matches: 246  
Percent Similarity: 96.15% Conservative: 4  
Best Local Similarity: 94.62% Mismatches: 9  
Query Match: 88.06% Indels: 3  
DB: Gaps: 0  
US-09-905-674-2 (1-270) x B1663945 (1-831)  
OY 1 MethIstrYrYrArGlyrSerAsnAlaIlysaIlySerCysTrpIlystrYrIleuIleuPhe 20  
Db 52 ATGCATATTTATGATCTCGAAGCCGAGGTGAGTGTGATACAAAGTACTGCTCTTC 111  
OY 21 SerTYrAsnIleIlePheTrpleuAlaGlyAlaIlyPheIleuGlyAlaGlyIleuTrpAla 40  
Db 112 AGCTACATATCGTCTTTGGCTGGCTGGAGTGTCTTCTTGAGAGTGGGCTGTGGCA 171  
OY 41 TrpserGluYsgIlyAlaIleuSerAspIleuThrIlyValThrArGmeThIsgIlyIleasp 60  
Db 172 TGGAGCGAAAGGGTGTGCTGTCCGACCTCACCAAGTGCACCGGTTCATGGAATCGAC 231  
OY 61 ProValIleuValIleuMetValGlyAlaIlyMetPheThrIleuGlyPheAlaGlyCys 80  
Db 232 CCGGTGCTGCTGCTGTGATGTGGGTGGTGTGATGTTCACACTGGGATTCGCAAGCGCTGT 291  
OY 81 ValGlyAlaIleuArGluAsnIleCysIleuIleuAsnIlePheCysGlyThrIleValIleu 100  
Db 292 GTGGGGGCCCTCCAGAGAACATGCTGCTCAAGTTTCTGTGGGGCCATTGTGCTC 351  
OY 101 IlePhePheIleuGlnIleuAlaIlyAlaIlyPheIleuPheGlnAspTrpValArg 120  
Db 352 ATCTTCTTCTGGAACGTGCGGTGCGGTGTGGCTTTTATTTCCAAAGCTGGGTGAGA 411  
OY 121 ASPATGPhaerGluPhepGluSerAsnIleYsserTYrArGAspAspIleAspIleu 140  
Db 412 GACCGGTTCGGGAATCTTCCAGAGCAACATCAAGTCTATCGGATGCATCGACCTG 471  
OY 141 GlnAsnIleAspSerIleuGlnIlysaAsnGlnCysGlyAlaTyrGlyProGlu 160  
Db 472 CAGAACCTCATGACTCCCTTCAGAAAGCTAATCAGTGTGCGGGGCTTACGGGCTCGAA 531  
OY 161 ASPPTPASPleuAsnValTYrPheAsnCYsserGlyAlaSerTYrSerArGluIlyscys 180  
Db 532 GACTGGACCTCAATGTCATCTTCACTGCAGTGTGCTGCACCTACACCGAAGAAATGT 591  
OY 181 GlyValProPheSerCysValProAspProAlaGlnIlyValAlaAsnThrGlnCys 200  
Db 592 GGGGACCTTCTCTGCTGTGTCAGATCTCCACAAAAAGTCGGAACACACACTGT 651  
OY 201 GLYTYrAspValArgIleGlnIleuIlysserIlystrPAspIuserIlePheThrIlysgly 220  
Db 652 GGCTATGATGTCGGATTCACAGTGAAGAGCG-GGGATGATGTCATCTTACAAAAAGGA 710  
OY 221 CysIleGlnAlaIleuGlnSerTrpleuProArGAsnIleTYr-IleValAlaGlyAlaIlyph 240  
Db 711 TGCAATCCAGGCTGTGGAAGGCTGCTGCCAGAACATCTACCAATTTGTGCTGTCTT 770  
OY 240 eIleAlaIleSerIleuGlnIlePheGlyIlePheIleuAlaArgThrIleuIleser 259  
Db 771 CATGCGATCTCACTGTGCAAGATTTTGGCATCTTCG-GGGAGGACCTGTATCTCG 827  
RESULT 9  
B0572112 742 bp mRNA linear EST 19-JUN-2002  
LOCUS B0572112  
DEFINITION UT-M-FD0-byd-n-13-0-UT.r1 NIH\_BMAP\_F00 Mus musculus cDNA clone



IMAGE:5716812 5', mRNA sequence.  
 ACCESSION B0572112  
 VERSION B0572112.1 GI:21475429  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 742)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Dr. James Lin, University of Iowa  
 CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA library arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILIN at: <http://image.llnl.gov>  
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
 Seq primer: pYX-5.  
 FEATURES  
 source Location/Qualifiers  
 1..742  
 /organism="Mus musculus"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5716812"  
 /clone\_1lb="NIH-BMAP\_FD0"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 12.5 dpc"  
 /lab\_host="DH10B (TI phage resistant)"  
 /note="Organ: brain; Vector: pYX-Asc; Site\_1: Ecor I; Site\_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAGAGGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemm Chin, Ph.D., program coordinator."  
 BASE COUNT 151 a 184 c 212 g 193 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,94e-146 Length: 742  
 Score: 1275.00 Matches: 235  
 Percent Similarity: 96.75% Conservaive: 3  
 Best Local Similarity: 95.53% Mismatches: 8  
 Query Match: 87.99% Indels: 0  
 DB: 14 Gaps: 0  
 US-09-905-674-2 (1-270) x B0572112 (1-742)  
 QY 1 Methistyrtyrargtyrsersanlalysvalsercysrtprtyltyrleuonphe 20  
 |||||||  
 DB 3 ATGCACATATTATGATCTCGAACCGCGAGGTGCTGTTACAACTGCTCTTC 62  
 |||||||  
 QY 21 SerTyrAsnIlellepheTrpleuAaglyValAlpheuGlyValGlyLeuTrpAla 40  
 |||||||  
 DB 63 AGCTACAATATGCTCTTTGGCTGGCTGAGATTGTCCTTCCTTGAGGTGGGCTG 122  
 |||||||  
 QY 41 TTPserGluLysGlyValleuSerAspLeuTrpLysValTharGMeHisGlyLeasp 60  
 |||||||  
 DB 123 TGGAGCAAAAGGtGTGCTGTCCGACCTCACCAAGGtGACCCGCTGCATGCAATGC 182  
 |||||||

QY 61 ProValValleuValleuMetValGlyValValMetphenTrleuGlypheAlaGlyCys 80  
 |||||||  
 DB 183 CCGCTGTGCTGTGTGTATGATGCTGGCTGGTGTATGTCACACTGGATGCGAGCTGT 242  
 |||||||  
 QY 81 ValGlyAlaLeuArgGlyLysnIlCysLeuLeuAsnPhenCysGlyTrpIleValLeu 100  
 |||||||  
 DB 243 GTGGGGGCCCTCCGAGNACATCTGCTGCTCAAGTTTCTGTGGGCCATTGTCTCT 302  
 |||||||  
 QY 101 IllephenLeuGluLeuAlaValAlaValLeuAlaPhenLeuPhenGlnAspTrpValArg 120  
 |||||||  
 DB 303 ATCTTCTTCGGAACCTGCGCGCTGGCTGTGGCTTTTATTCGAAGACTGGGTGAGA 362  
 |||||||  
 QY 121 AsparGpheArgGluPhenPheGluSerAsnIlleYserTyrArgAspPheLeu 140  
 |||||||  
 DB 363 GACCGTTCCGGAAATTTCTGAGAGCAACATCAAGTCTATCGGATGACATCGACTTG 422  
 |||||||  
 QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTrpGlyProGlu 160  
 |||||||  
 DB 423 CAGAACTCATTTGACTCCTTCAGAAAGCTATACAGTGTGCGGGCTTAGCGCCCTGAA 482  
 |||||||  
 QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180  
 |||||||  
 DB 483 GACTGGGACCTCAATGCTCTACTCACTGACGTGTGCTCAGCTACAGCCGAGAAATGT 542  
 |||||||  
 QY 181 GlyValIProPheSerCysValProAspProAlaGlnLysValValAsnThrGlnCys 200  
 |||||||  
 DB 543 GGGGTACCTTCTCTCTGCTGTGTGCTCCAGATCTCTCACANAAAGTCTGGAACACACACTGT 602  
 |||||||  
 QY 201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheTrpLysGly 220  
 |||||||  
 DB 603 GGCTATGATGTCCGAGATTCACTGAGAGCAAGTGGATGAGTTCATCTWTTCAAAGGA 662  
 |||||||  
 QY 221 CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240  
 |||||||  
 DB 663 TGCAATCCAGGCTGTGGAAGGCTGCTGCCAGCAACATCTACATGTGCTGTCTTC 722  
 |||||||  
 QY 241 IleAlaIleSerLeuLeu 246  
 |||  
 DB 723 ATTTCATCTCTACTGCTG 740  
 |||  
 RESULT 10  
 AL519790 951 bp mRNA linear EST 13-FEB-2001  
 LOCUS  
 DEFINITION AL519790 tLTL NFE004\_NBC2 Homo sapiens CDNA clone CS0DB0051602 5  
 prime, mRNA sequence.  
 ACCESSION AL519790  
 VERSION AL519790.1 GI:12783283  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 951)  
 AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr).  
 FEATURES  
 source Location/Qualifiers  
 1..951  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DB0051602"  
 /clone\_1lb="LTL\_NFE004\_NBC2"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and





```
|||||
Db 367 GGGGACCGGTTCCGGGAGTCTTCGAGAGCAACATCAAGTCTACCGGGAGCATATCGATC 426
Qy 140 euglnasnleuileaspserrleuglnlysalaasngincyscysgylvalatyrglyprog 160
Db 427 TGGAAAACCTCATCGACTCCCTTCAGAAAGCTAACCAAGTCTGTGTGGCGCATATGAGCCCTG 486
Qy 160 lualsprrpaspleuasnvaltyrphasnscysserglyalaserlyrseratrglylasc 180
Db 487 AAGCTGAGGACCTCAACGCTTCAATTCAGCGGCTGCGACCTACAGCGGAGAGAACT 546
Qy 180 ysglyvalpropheserCysCysValProaspProalaglnlysalvalasnthglnc 200
Db 547 GCGGGGTCCCTCTCTCTGCTGCTGCGCATCTCCGCAAAAAAGTCGTAACACACAGT 606
Qy 200 ysgly-Tyr-AspValArgIleGlnleuIlySerIlySTPaspGluSerIlephenhrly 219
Db 607 GTGGCATATCATGTCAGATTTAGCTGAGAGCAAGTGGATGAGTCCATCTTCACGAA 666
Qy 219 sglcysIleGlnAlaLeuGln-SerTrpleu-ProArgAsnIleTyrlle-ValAlaGl 238
Db 667 AGCTCATCTCCAGCGCTGGAAACGCTGCTCCCGCGAACAATTACATTTGGAGCTGG 726
Qy 238 yValPheIleAlaIleSerleuLeu---GlnIlePheGlyIlePheLeuAlaArgThr-L 257
Db 727 CGTCTTCATGCGCATCTCGCTGGTTCGAAATATTTGGGATCTTCTGCGCAAGAGCGCC 786
Qy 257 euileaserasp 260
Db 787 TGATCTCAGAC 797
RESULT 12
BQ948800 870 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT.8806962 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6338432
DEFINITION 5', mRNA sequence.
ACCESSION BQ948800
VERSION BQ948800.1 GI:22364278
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 870)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@phs.fda.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L12CM2531 row: k column: 09
High quality sequence start: 3
High quality sequence stop: 530.
Location/Qualifiers
1..870
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6338432"
/clone_id="NIH_MGC_47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pOTB; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
```

```
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."
BASE COUNT 175 a 223 c 257 g 214 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 4,04e-142 Length: 870
Score: 1246.50 Matches: 241
Percent Similarity: 94.51% Conservative: 0
Best Local Similarity: 94.51% Mismatches: 9
Query Match: 86.02% Indels: 5
Db: 14 Gaps: 2
US-09-905-674-2 (1-270) x BQ948800 (1-870)
Qy 3 TyrTyraGtyrSerAsnAlaIysValSerCysTrpTylySTyrlleuLeuPheSerTyr 22
Db 4 TATTATAGATACCTCAAGCCAGGTCAAGTGTGTAACAAGTACCTCTTCAGCTAC 63
Qy 23 AsnIleIlePheTrpleuAlaGlyValValPheLeuGlyValIglyleuTPalaTrpSer 42
Db 64 AACATCATCTTCTGGTGGCTGGAGTGTCTTCTTGAGTCGGGCTGGGCAATGGAGC 123
Qy 43 GluysglyValIleuSerAspLeuThrIlyValThrArgmetHlsgIlyleasProVal 62
Db 124 GAAAGAGGTGTGCTGTGCGACCTCACCACCAAGTACCCGGATGCTAGTAATCGACCTGTG 183
Qy 63 ValIleuValIleuMetValGlyValValMetPheThrleuGlyPheAlaGlyCysValGly 82
Db 184 GTGCTGTCTCGATGATGGGCGGTGGTGTGATGTCACCTGGGGGTTCGCGGCTGGTGGGG 243
Qy 83 AlaIleuArgGluAsnIleCysLeuLeuAsnPhetheCysGlyThrIleValIlePhe 102
Db 244 GCTCTGGGGAGATATCTGTGCTCAACTTTTCTGTGGACCACTCGCTCATCTTC 303
Qy 103 PheLeuGlnleuAlaValAlaValIleuAlaPheLeuPheGlnAspTrpValArgAspArg 122
Db 304 TTCCTGGAGCTGGCTGTGGCGGTGCTGCTCTCTCTTCAGGACATGGCGTGAAGACCG 363
Qy 123 PheArgGluPhePheGluSerAsnIleIysSerTyrArgAspAspIleaspleuGlnAsn 142
Db 364 TTCGGGAGTCTTCGAGACCAACATCACTACCGGAGCATATTCATCTGCAAAAC 423
Qy 143 LeuIleAspSerleuGlnIysAlaAsnGlnCysCysGlyAlaTyrglyProGluAspTrp 162
Db 424 CTGATTCAGATCCCTTCGAAAGACTAACCAAGTGTGTGGCCATATGGCCCTGAAGACTGG 483
Qy 163 AspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluIysCysGlyVal 182
Db 484 GACCTCAACGCTCTCAATTCATTTGAGCGGTGCCAGTACAGCGGAGAAAGTGGGGTTC 543
Qy 183 ProPheSerCysCysValProaspProalaglnlysalvalasnthrgincysgyltyr 202
Db 544 CCTTCTCTGCTGCGCGGCGAGATCTCGCAAAAGTTGTGAACACACAGTGTGATAT 603
Qy 203 AspValArgIleGlnleuIlySerIlySTPaspGluSerIlephenhrlysglyCysIle 222
Db 604 GATTCATGATTCATGACGAGAGCAAGTGGATGATCTTCACGAAAGGCTCATC 663
Qy 223 GlnAlaLeuGlnSerTrp-LeuProArg-AsnIleTyrlleValAla---GlyValPheI 241
Db 664 CAGGCGGTGGAAGAGTGGGCTCCCGGAGAACATTTACATTTGGGGTGGCGTCTCTCA 723
Qy 241 leuAlaIleSer-----LeuLeuGlnIlePheGlyIlePhe 252
Db 724 TTCGCCATCTCCGCTGTGTCANAAATATTTTGGCATCTTT 764
RESULT 13
BG765719 755 bp mRNA linear EST 15-MAY-2001
LOCUS BG765719
DEFINITION 602739738F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4869633 5',
mRNA sequence.
```

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
BG765719	1	GI:14076372	EST	human.					
ORGANISM				Homo sapiens					
REFERENCE				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS				Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE				NIH-MGC http://mgc.nci.nih.gov/.					
JOURNAL				National Institutes of Health, Mammalian Gene Collection (MGC)					
COMMENT				Unpublished (1999)					
				Contact: Robert Strausberg, Ph.D.					
				Email: c9apbs-remail.nih.gov					
				Tissue Procurement: ATCC/DOCD/DPH					
				cDNA Library Preparation: Ling Hong/Rubin Laboratory					
				cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)					
				DNA Sequencing by: Incyte Genomics, Inc.					
				Clone distribution: MGC clone distribution information can be					
				found through the I.M.A.G.E. Consortium/LMNL at:					
				http://image.lnl.gov					
				Plate: L16CML741 row: k column: 10					
				High quality sequence stop: 754.					
FEATURES				Location/Qualifiers					
SOURCE				1..755					
				/organism="Homo sapiens"					
				/db_xref="taxon:9606"					
				/clone="IMAGE:4869633"					
				/clone_lib="NIH-MGC_49"					
				/tissue_type="melanotic melanoma, high MDR (cell line)"					
				/lab_note="DH10B (phage-resistant)"					
				/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:					
				EcoRI; cDNA made by oligo-dT priming. Directionally cloned					
				into EcoRI/XhoI sites using the following 5' adaptor:					
				GGCGCAG(G). Size-selected >500bp for average insert size					
				1.8kb. Library constructed by Ling Hong in the laboratory					
				of Gerald M. Rubin (University of California, Berkeley)					
				using Zap-cDNA synthesis kit (Stratagene) and Superscript					
				II RT (Life Technologies). Note: this is a NIH-MGC					
				Library."					
BASE COUNT				143 a 207 c 226 g 179 t					
ORIGIN				Library."					
Alignment Scores:									
Pred. No.:				6.53e-142					
Score:				1244.00					
Percent Similarity:				100.00%					
Best Local Similarity:				99.56%					
Query Match:				85.85%					
DB:				12					
				gaps:					
				0					
US-09-905-674-2 (1-270) x BG765719 (1-755)									
QY 1	Methistyrtyrargtyrserasnalalysvalsercystyrtyrtyrleu	Phe	20						
Db									
68	ATGCACATATATAGATACTCTAAGCCAAAGTCAGCTGCTGTACAGTACTCTTTC	127							
QY 21	SertryrsnleilepethrpreuAlaGlyValValPheleuGlyValGlyLeu	TrpAla	40						
Db									
128	AGCTACACATCATCTTCTGCTGCTGCGAGTGTCTTCTTCTTGAAGTCGGGCTGTGGCA	187							
QY 41	TrpSerGluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGly	Leuasp	60						
Db									
188	TGGAGCGAAAGGCTGTGCTGCTCCGACCTCACCAAG								

	BASE COUNT	153 a	216 c	243 g	197 t	1 others
	ORIGIN					
Dn	368	ATCTTCTCCGAGACGTGGCTGTGGCGTGCTGCCTTCCTCTTTCCAGAGCATGGTAGAG	427			
Oy	121	Asparagpheargslunherhegluserashllylssettyrargaspshlleaspleu	140			
Dd	428	GACGGGTCCGGAGATTCTTCGAGACAACATCAACTCCTACC GGACGATTCGATCTG	487			
Oy	141	Glhasleu1leaasperseuglnlysaliaasnlnlcyscsglyalalyrglyprrogu	160			
Dd	488	CAAAACCTATGACATCCCTTCAGAAAGCTAACCACTGCTGTGGCCATATGGCCTCGAA	547			
Oy	161	AseptaspheuasnaValTyrrheaenCysSerglyAlaserTyzserArglnulyCys	180			
Dd	548	GACTGGGACCTCAACACTTACTTCAATTCAGAGGGTGGCCAGGTACAGCCGAGAAGTGC	607			
Oy	181	glyvalProphsercscysvalProappproalaglInlyvalYalsnthrgIncyS	200			
Dd	608	GGGGTCCCTTCTCCGTCGGCGGCGAGTCTGCCCAAAGTCTGTGAACACACAGTGT	667			
Oy	201	gLyTyRASPvAlArllleglnleuIlySSerlySTPRspgluserlllepherThrylgly	220			
Dd	668	GGATATGATGTCAGGATTCAGCTGGAAGACCACTGGAGTGCATCTTCAGAAAGGT	727			
Oy	221	CysllleglnAlaleuGlusertPleu	229			
Dd	728	TGCATCCAGGCGCTGAAAAAGCTGGTGC	754			
RESULT 14	BG760633	810 bp	mRNA	linear	EST 15-MAY-2001	
LOCUS	BG760633	602717387/1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4840916.5,				
DEFINITION	mRNA sequence.					
ACCESSION	BG760633					
VERSION	BG760633.1 GI:14071286					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE	1 (bases 1 to 810)					
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D. Email: rs9ab@remail.nih.gov Tissue Procurement: ATCC/DCTP/DP CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC Clone distribution information can be found through the I.M.A.G.E. Consortium/LMLN at: <a href="http://image.lnl.nih.gov">http://image.lnl.nih.gov</a> Plate: LILCM1672 row: n column: 21 High quality sequence stop: 782.					
FEATURES	Location/Qualifiers					
SOURCE	1..810					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone="IMAGE:4840916"					
	/clone_lib="NIH_MGC_49"					
	/tissue_type="melanotic melanoma, high MDR (cell line)"					
	/lab_host="DH10B (phage-resistant)"					
	/note="Organ: skin; Vector: pOTB; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."					

Alignment Scores:

Pred. No.:	6,72e-140	Length:	810
Score:	1228.00	Matches:	238
Percent Similarity:	97.54%	Conservative:	0
Best Local Similarity:	97.54%	Mismatches:	6
Query Match:	84.75%	Indels:	2
DB:	12	Gaps:	0

US-09-905-674-2 (1-270) x BG74843 (1-810)

```

Oy 1 MethISTYTYRARGTYRSerAsnAlaIysValSerCysTrpTYRlyLysTYrLeuLeuPhe 20
Db 68 ATGCACATATATAGTACTCTTAACCCCAAGGTCACTGCTGTGACAGTACTCTCTTTC 127
Oy 21 SerTYrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAla 40
Db 128 AGCTACAAACATCATCTTCTGCTGAGTGGCTGAGTCTTCTTGTGAGAGTCGGCTGGGCA 187
Oy 41 TrpSerGluIysGlyValLeuSerAspLeuThrIysValThrArgMetHisGlyIleAsp 60
Db 188 TGGAGCGAAAGGGTGTCTGTCACCTCACAAAGTACCCGAGTGCATGAAATCGAC 247
Oy 61 ProValIleLeuValIleMetValGlyValIleMetPheThrLeuGlyPheAlaGlyCys 80
Db 248 CCTGTGCTGCTGCTCTGATGCTGGGCGTGGTGATGTTCAACCTGGGCTTGGCGGCTGC 307
Oy 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
Db 308 GTGGGGGCTCTCGGAGAAATATCTGTGCTGCACACTTTTCTGTGGCACCATGCTGCTC 367
Oy 101 IlePhePheLeuGluLeuAlaValAlaValIleuAlaPheLeuPheGluAspTrpValArg 120
Db 368 ATCTTCTTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
Oy 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTYrArgAspPheIleAspLeu 140
Db 428 GACCGGTCGGGAGTCTTCTTCAGAGCAACATCAAGTCTTACCGGAGCATGATCATCTG 487
Oy 141 GluAsnIleuIleAspSerLeuGluIleLysAlaAsnGlnCysGlyValIleTYrGlyProIu 160
Db 488 CAAACCTCATGACATCCCTTCAGAAAGCTAACAGCTGTGGGCAATAGNGCTGAA 547
Oy 161 AspTrpAspLeuAsnValTYrPheAsnCysSerGlyAlaSerTYrSerArgGluIysCys 180
Db 548 GACTGGGACCTCAAGCTCTACTTCAATTGCACGGTGCAGGTACAGCGAGAGAGAGTC 607
Oy 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValIleAsnThrGlnCys 200
Db 608 GGGGTCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667
Oy 201 GlyTYrAspValArgIleIleuLysSerIysTrpAspGluSerIlePheThrIysGly 220
Db 668 GGATTAAGATGATGAGTTAGTGAAGAGCAAGTGGATGATGATGATGATGATGATGATG 727
Oy 221 CysIleGluAlaLeuGluSerTrpLeuProArgAsnIleTYrIleValAlaGlyValPhe 240
Db 728 TG-ATCCAGGCGCTGAAAGGCTGGGTCCGGGCAACATTATATGTGTGT-GGCTCTTC 785
Oy 241 IleAlaIleSer 244
Db 786 ATCCGATCTCG 797

```

RESULT 15  
BG974843 770 bp mRNA linear EST 12-JUN-2001  
LOCUS 602842725F1 NCI\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:4978446 5',  
DEFINITION mRNA sequence.  
ACCESSION BG974843  
VERSION BG974843.1 GI:14362480  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS NIH-MGC  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgrabs-ri@mail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch  
Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL0975 row: a column: 07  
High quality sequence stop: 760.  
Location/Qualifiers

FEATURES  
source 1..770  
/organism="Mus musculus"  
/strain="NMRI"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4978446"  
/clone\_lib="NCI\_CGAP\_Mam4"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; Salt: Site:2; NotI: Cloned unidirectionally. Primer: Oligo dt. Library constructed by: Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furch. NIH Reference for transgenic model: Li et al., Cell growth and differentiation 7, 3-11 (1996)."

BASE COUNT  
152 a 226 g 192 t  
ORIGIN

Alignment Scores:

Pred. No.:	2.55e-139	Length:	770
Score:	1223.00 <td>Matches:</td> <td>228 </td>	Matches:	228
Percent Similarity:	97.06% <td>Conservative:</td> <td>3 </td>	Conservative:	3
Best Local Similarity:	95.80% <td>Mismatches:</td> <td>5 </td>	Mismatches:	5
Query Match:	84.40% <td>Indels:</td> <td>2 </td>	Indels:	2
DB:	13	Gaps:	0

US-09-905-674-2 (1-270) x BG974843 (1-770)

```

Oy 1 MethISTYTYRARGTYRSerAsnAlaIysValSerCysTrpTYRlyLysTYrLeuLeuPhe 20
Db 53 ATGCACATATATAGTACTCTTAACCCCAAGGTCACTGCTGTGACAGTACTCTCTTTC 112
Oy 21 SerTYrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAla 40
Db 113 AGCTACATATATAGTACTCTTGGCTGCTGAGTGTCTTCTTGTGAGTGGGCGTGGGCA 172
Oy 41 TrpSerGluIysGlyValLeuSerAspLeuThrIysValThrArgMetHisGlyIleAsp 60
Db 173 TGGAGCGAAAGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 232
Oy 61 ProValIleLeuValIleMetValGlyValIleMetPheThrLeuGlyPheAlaGlyCys 80
Db 233 CCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 292
Oy 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
Db 293 GTGGGGGCGCTCCAGAGAAATATCTGCTGCTCAAGTTTCTGTGGGCGCATTTGCTCTC 352
Oy 101 IlePhePheLeuGluLeuAlaValAlaValIleuAlaPheLeuPheGluAspTrpValArg 120
Db 353 ATCTTCTTCTGAGAACTGGCGCTGCTGCTGCTGCTGCTTATTATCCAAAGACTGGTGAGA 412
Oy 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTYrArgAspPheIleAspLeu 140

```

```

Db      413 GACCGGTTCCGGGAATCTTCGAGACAACATCAAGTCCTATCGGGATGACATCGACCCTG 472
QY      141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGlu 160
         |||||
Db      473 CAGAACTCATTTGACTCCCTTCAGAAAGCTAATACAGTGTGGGGCTTACGGCCCTGAA 532
QY      161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180
         |||||
Db      533 GACTGGAGACCTCAATGTCTACTTCAACTGCAAGTGTGGCCAGCTACAGCCGAGAGAAATGT 592
QY      181 GlyValProPheSerCysCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
         |||||
Db      593 GGGGTACCCCTTCTCCTGCTGTGTGCCAGATCCTGCACAAAAGTCGTGACACACACAGTGT 652
QY      201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220
         |||||
Db      653 GGCTATGATGTCCGATTCAGCTGAAGAGCAAGTGGGATGAGTTCATCTTTACAAAAGGA 712
QY      221 CysIleGln-AlaLeuGluSerTrpLeuProArg-AsnIleTyrIleVal 236
         |||||
Db      713 TGCATCCAAAGGCTCTGGAAGGCTGGCTGCCCGAAGACATCTACATTGTG 762

```

Search completed: December 19, 2002, 07:45:49  
 Job time : 2012 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 19, 2002, 06:19:07 ; Search time 51 Seconds

(without alignments)  
1623.583 Million cell updates/sec

Title: US-09-905-674-2

Perfect score: 1449

Sequence: 1 MHYRYSNMAVSCWYKYLFL.....IFLARLISDIEAVKGGHMF 270

## Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Xgapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+lp2n.model -DEV=xlh  
-o=/cgn2\_1/USPTO/spool/US0905674/runat\_16122002\_111846\_6791/app\_query.fasta.1.455  
-DB=Issued.patents.NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=2000000000  
-USER=US0905674\_ECGN1.1.17\_etrunat\_16122002\_111846\_6791 -NCPU=6 -ICPU=3  
-NO\_XLPTX -NO\_MAP -LARGEXERT -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued.patents.NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCtUS.COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	278	19.2	1151	2	US-08-807-044-2
2	266	18.4	827	4	US-09-333-599-5
3	266	18.4	870	4	US-09-333-599-1
4	264.5	18.3	1344	3	US-08-705-771-8
5	259.5	17.9	1452	2	US-08-807-044-4
6	259.5	17.9	1452	5	PCR-US91-04986-1
7	257.5	17.8	687	1	US-08-254-493-2
8	257.5	17.8	687	1	US-08-408-222B-2
9	257.5	17.8	1120	1	US-08-408-222B-3
10	257.5	17.8	1624	4	US-08-430-225A-19
11	252.5	17.4	977	2	US-08-855-140-2
12	252.5	17.4	1120	1	US-08-254-493-3

13	218.5	15.1	933	3	US-08-808-148-2	Sequence 2, Appl1
14	218.5	15.1	1289	4	US-09-020-956-111	Sequence 111, App
15	218.5	15.1	1289	4	US-09-030-607-111	Sequence 111, App
16	218.5	15.1	1289	4	US-09-605-785-111	Sequence 111, App
17	218.5	15.1	1289	4	US-09-439-313-111	Sequence 111, App
18	218.5	15.1	1289	4	US-09-352-616A-111	Sequence 111, App
19	218.5	15.1	1289	4	US-09-232-149A-111	Sequence 111, App
20	198.5	13.7	1001	3	US-08-705-771-6	Sequence 6, Appl1
21	197.5	13.6	399	3	US-09-188-930-62	Sequence 62, App
22	197.5	13.6	399	3	US-09-188-930-243	Sequence 243, App
23	176.5	12.2	801	4	US-09-020-956-16	Sequence 16, Appl
24	176.5	12.2	801	4	US-09-030-607-16	Sequence 16, Appl
25	176.5	12.2	801	4	US-09-605-785-16	Sequence 16, Appl
26	176.5	12.2	801	4	US-09-439-313-16	Sequence 16, Appl
27	176.5	12.2	801	4	US-09-352-616A-16	Sequence 16, Appl
28	176.5	12.2	801	4	US-09-232-149A-16	Sequence 16, Appl
29	164	11.3	398	4	US-09-385-982-273	Sequence 273, App
30	161.5	11.1	560	4	US-09-221-298-12	Sequence 12, Appl
31	152.5	10.5	740	4	US-09-020-956-17	Sequence 17, Appl
32	152.5	10.5	740	4	US-09-030-607-17	Sequence 17, Appl
33	152.5	10.5	740	4	US-09-605-785-17	Sequence 17, Appl
34	152.5	10.5	740	4	US-09-439-313-17	Sequence 17, Appl
35	152.5	10.5	740	4	US-09-352-616A-17	Sequence 17, Appl
36	152.5	10.5	740	4	US-09-232-149A-17	Sequence 17, Appl
37	147.5	10.2	1966	4	US-08-957-130-14	Sequence 14, Appl
38	141.5	9.8	1782	4	US-09-149-476-120	Sequence 120, Appl
39	136.5	9.4	751	4	US-09-020-956-12	Sequence 12, Appl
40	136.5	9.4	751	4	US-09-030-607-12	Sequence 12, Appl
41	136.5	9.4	751	4	US-09-605-785-12	Sequence 12, Appl
42	136.5	9.4	751	4	US-09-439-313-12	Sequence 12, Appl
43	136.5	9.4	751	4	US-09-352-616A-12	Sequence 12, Appl
44	136.5	9.4	751	4	US-09-232-149A-12	Sequence 12, Appl
45	133	9.2	729	4	US-09-020-956-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-08-807-044-2  
Sequence 2, Application US/08807044  
Patent No. 5863735  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
TITLE OF INVENTION: NOVEL HUMAN TRANSMEMBRANE 4 SUPERFAMILY  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTED for windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/807, 044  
FILING DATE: Herewith  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0224 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1151 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SCORNOT01  
CLONE: 663655  
US-08-807-044-2

## Alignment Scores:

Pred. No.:	3,32e-25	length:	1151
Score:	278.00	Matches:	75
Percent Similarity:	44.688	Conservative:	51
Best Local Similarity:	26.608	Mismatches:	94
Query Match:	19.198	Indels:	62
DB:	2	Gaps:	13

US-09-905-674-2 (1-270) x US-08-807-044-2 (1-1151)

```
QY 16 LysTyrLeuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGly 35
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 AAGTACTCATGTTGCTGCTTCAACGTCGTCCTGCTGGAGGCTGCGCTGCTGGCT 300

QY 36 ValGlyLeuTrpAlaTrpSerGlyValLeuSerAspLeuThrLysValThrArg 55
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 GTGCGCATCTGCTGGCCGACACAGGGG-----AGCTTGCCACGCTGCTCTTCC 354

QY 56 MetHisGlyIleAspProValValLeuValLeuMetValGlyValValMetThrIleu 75
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 355 TTCCCGTCCCTGCTGCTGCCAACCTGCTCATCATCATCCGCGCTTGTGCTGGCCATC 414

QY 76 GlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnIlePheCys 95
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 415 GCGTCTGCTGGCTGCTGGGCTGCCATCAAGAACAGTGCCTGCTGCTGCTGCTTCTTC 474

QY 96 GlyThrIleValIlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPhe 115
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 475 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534

QY 116 GluAspTrpValArgAspTrpPhe--ArgGluPhePheGluSerAsnIleLysSerTyr 134
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 535 ACGGACAAAGATT--GACAGGATGCGCCAGCAAGACCTGTAAGAAAGCTTGCACTGTAC 591

QY 135 -----ArgAspAspIleAspLeuGluAsnLeuIleAspSerLeuGluLysAlaAsn--- 151
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 592 GGCACGACGAGGCAAGCTGGGCTCAACACAGCGCTGAGCATCATCAAGACGACCTCGA 651

QY 151 ----- 151
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 652 GCGGTGGGAGGTGGGCGGGTGGGCGGCTGCCCTCCCTNCTGCTCAACCCGACCT 711

QY 152 -----GlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsn 165
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 712 GAGCTTGCCCCCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768

QY 166 ValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCysGlyValProPheSer 185
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 769 GTGTAC-----AACGCCAGCGGGGTACTGACTGCTC 798

QY 186 CysGlyValProAspProAlaGluLysValValAsnThrGlnCysGlyTyrAspValArg 205
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 799 TCCTGCTTGAG-----TTCACTGAGAGCTGTGGG----- 828

QY 206 IleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGlnAlaLeu 225
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 829 CTGCACGCCCCCGGACCTGG-----TGAAGCGCCGCTGTACAGACGGGTG 876

QY 226 GluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe---IleAlaIleSer 244
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 877 AAGGTGTGCTTCAGAGAACCTG---CTGGCTGTGGGCTGTGGGCTGTGCGACGGCG 933
```

```
QY 245 LeuLeuGlnIlePheGlyIlePheLeuAlaArgThrIleLeuSerAspIleGluAlaVal 264
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 934 CTGTGCAGATCTGCGGCTGACCTTCCGATGACCATGTACTGCAAGTG-----GTC 987

QY 265 LysAla 266
|||||
Db 988 AAGGCA 993
```

## RESULT 2

US-09-333-599-5  
Sequence 5, Application US/09333599  
Patent No. 6245898

## GENERAL INFORMATION:

APPLICANT: Testa, Jacqueline E.  
APPLICANT: Quigley, James P.  
APPLICANT: Seandel, Marco  
TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS  
TITLE OF INVENTION: ASSOCIATED WITH TUMOR METASTASIS  
FILE REFERENCE: SONY  
CURRENT APPLICATION NUMBER: US/09/333,599  
CURRENT FILING DATE: 1999-06-15  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: patentIn Ver. 2.0

SEQ ID NO 5  
LENGTH: 827  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (57)..(815)  
US-09-333-599-5

## Alignment Scores:

Pred. No.:	6.38e-24	length:	827
Score:	266.00	Matches:	74
Percent Similarity:	44.488	Conservative:	51
Best Local Similarity:	26.338	Mismatches:	108
Query Match:	18.368	Indels:	48
DB:	4	Gaps:	11

US-09-905-674-2 (1-270) x US-09-333-599-5 (1-827)

```
QY 6 TyrSerAsnAlaLysValSerCys-----TrpTyrLysTyrLeuLeuPheSerTyr 22
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 66 TTCAACGAGAGAGAACATGATGCGCACCGTTTGCTCAAGATGACGCTGTATACCTAC 125

QY 23 AsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAlaTrpSer 42
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 126 AATTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185

QY 43 GluLysGly--ValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspPro 61
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 186 CTCAAGAGTGAATCATCATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 236

QY 62 ValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysVal 81
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 237 GCTTACATCTCGTGGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 296

QY 82 GlyAlaLeuArgGluAsnIleCysLeuLeuAsnIlePheCysGlyThrIleValLeuIle 101
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 297 GCCACCTTCAAGAGGCTGCGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 356

QY 102 PhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArgAsp 121
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 357 TTTCTGCTGAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 416

QY 122 ArgPheArgGluPhePheGluSerAsnIle--LysSerTyrArgAsp-----AspIle 138
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 417 GAGCTCAAGAGGAACCTGAGACACCATGACCAAGCGGTACACACGCGGCGCATGAG 476

QY 139 AspLeuGlnAsnLeuIleAspSerLeuGlnLysAlaAsnIleCysGlyAlaTyrGly 158
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 477 GCTGTGACCAAGCGCTGAGACAGCTGACAGGAGTTCCACTGCTGTGCGACCAACAC 536
```







Db 203 -----TTCCAAACCTCCCTCCCTCACGCTGGGCAATGTTGTCTCATCTGGGCTCT 256  
Qy 71 ValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeu 90  
Db 257 ATATATCATGTAAGTTCCTCCCTCCGCGCTCATGAGGAGGCTCTATCAAGAAACAAGTGTCTG 316  
Qy 91 LeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaVal 110  
Db 317 CTATATGCTGTTCTTCATCCTGCGCTGATATTCCTCTGCTGAGAGGTGACCTGGCCATC 376  
Qy 111 LeuAlaPheLeuPheGlnAspTrpValAlaArgAspArgPheArgGluPhePheGluSerAsn 130  
Db 377 CTGCTCTTTGTATATGAACAAGAGGAATGATATGCTGCTAAGGCTCTGACCGACAGC 436  
Qy 131 IleLysSerTyrArgAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAla 150  
Db 437 ATCCACCGTTATCCACCTACACAAATACACCAAGCAGCGCTGGAGCTCCATCCAGTCATTT 496  
Qy 151 AsnGlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCys 170  
Db 497 CTGCAGTGTGTGTATTAATGCGACAGATGATGG----- 532  
Qy 171 SerGlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCysCysValProAsp 190  
Db 533 -----ACCACTGGCCCAACGACATCTTGC----- 556  
Qy 191 ProAlaGlnLysValAlaAsnThrGlnCysGlyTyrAspValArgIleGlnLeuLysSer 210  
Db 557 CCCTCAGATCGAAAGT----- 574  
Qy 211 LysTrpAspLysSerIlePheThrLysGlyCysIleGlnAlaLeuGluSerTrpLeuPro 230  
Db 575 -----GAGGCTGCTATGCAAGCAAGCAAGACTGTGGTTTAC 610  
Qy 231 ArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIlePheGly 250  
Db 611 TCCATATTCCTGTATATCGGAATCATCAACCATCTGTGTATGTGTATGAGGTGTGGGG 670  
Qy 251 IlePheLeuAlaArgThrLeuIleSerAspIleGlu 262  
Db 671 ATGCTCTTGCACCTGACCTCGAATCGCCAGATTGAC 706

RESULT 6  
PCT-US91-04986-1  
; Sequence 1, Application PC/TUS9104986  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian  
; APPLICANT: Allen, Janet  
; APPLICANT: Aruffo, Alejandro  
; APPLICANT: Camerini, David  
; APPLICANT: Lauffer, Leander  
; APPLICANT: Oquendo, Carmen  
; APPLICANT: Simmons, David L.  
; APPLICANT: Stamenkovic, Ivan  
; APPLICANT: Stengelin, Siegfried  
; APPLICANT: Amiot, Marlene  
; TITLE OF INVENTION: Rapid Immunoselection Cloning Method  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee & Associates  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/04986  
; FILING DATE: 19910715

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/553,759  
; FILING DATE: 13-JUL-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/498,809  
; FILING DATE: 23-MAR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/379,076  
; FILING DATE: 13-JUL-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/160,416  
; FILING DATE: 25-FEB-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wall, Margaret M.  
; REGISTRATION NUMBER: 33,462  
; REFERENCE/DOCKET NUMBER: 11-88C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303-499-8080  
; TELEFAX: 303-499-8089  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1452 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 74..733  
; PCT-US91-04986-1

Alignment Scores:  
Pred. No.: 9.83e-23 Length: 1452  
Score: 259.50 Matches: 62  
Percent Similarity: 42.86% Conservative: 46  
Best Local Similarity: 24.60% Mismatches: 89  
Query Match: 17.91% Indels: 55  
DB: 5 Gaps: 6

US-09-905-674-2 (1-270) x PCT-US91-04986-1 (1-1452)

Qy 16 LysTrpLeuLeuPheSerTyrAsnIleIlePheThrLeuAlaGlyValAlaPheLeuGly 35  
Db 101 AAGTATGCTCCGTGTTTCTTCAACTGCTCTTTGGAGTCTGTGCTGCTGATTTTGGGC 160  
Qy 36 ValGlyLeuTrp-----AlaTrpSerGluLysGlyValLeuSerAspLeuTrpLysVal 53  
Db 161 TTTGGGATCTACCTGCTGATCCACACAACTTCGAGTGTCTC----- 202  
Qy 54 ThrArgMetHisGlyIleAspProValValLeu-----ValLeuMetValGlyVal 70  
Db 203 -----TTCCATTAACCTCCCTCCCTCCACGCTGGGCAATGTTGTTCATCTGGGCTCT 256  
Qy 71 ValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeu 90  
Db 257 ATATATCATGTAAGTTCCTCCCTCCGCGCTCATGAGGAGGCTCTATCAAGAAACAAGTGTCTG 316  
Qy 91 LeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaVal 110  
Db 317 CTATATGCTGTTCTTCATCCTGCGCTGATATTCCTCTGCTGAGAGGTGACCTGGCCATC 376  
Qy 111 LeuAlaPheLeuPheGlnAspTrpValAlaArgAspArgPheArgGluPhePheGluSerAsn 130  
Db 377 CTGCTCTTTGTATATGAACAAGAGGAATGATATGCTGCTAAGGCTCTGACCGACAGC 436  
Qy 131 IleLysSerTyrArgAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAla 150  
Db 437 ATCCACCGTTATCCACCTACACAAATACACCAAGCAGCGCTGGAGCTCCATCCAGTCATTT 496  
Qy 151 AsnGlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCys 170  
Db 497 CTGCAGTGTGTGTATTAATGCGACAGATGATGG----- 532

QY 171 SerGIYAlaSerTYrSerArgIuIyScysGIYValProPheSerCysCysValProAsp 190  
| | | | |  
Db 533 -----ACCACTGGCCCAACGACATCTTGC----- 556  
QY 191 ProAlaGlnIySValAsnThrGlnCysGIYTYrAspValArgIleGlnLeuIySser 210  
| | | | |  
Db 557 CCTCAGATCGAAAGCTG----- 574  
QY 211 LysTrpAspGluSerIlePheThrLysGIYCysIleGlnAlaLeuGluSerTrpLeuPro 230  
| | | | |  
Db 575 -----GAGGGTGGCTATGCCAAGCAAGACACTGGGTTTCAT 610  
QY 231 ArgAsnIleTYrIleValAlaGlyValPheIleAlaIleSerLeuGlnIlePheGly 250  
| | | | |  
Db 611 TCCAATTTCCTGTATATCGAATCATCACCATCTGTGTATGTGTATGAGGTGTTGGGG 670  
QY 251 IlePheLeuAlaArgThrLeuIleSerAspIleGlu 262  
| | | | |  
Db 671 ATGTCCTTGCACTGACCTGAACTGCCAGATTGAC 706

RESULT 7  
US-08-254-493-2  
; Sequence 2, Application US/08254493  
; Patent No. 5439886  
; GENERAL INFORMATION:  
; APPLICANT: IKEYAMA, SHUICHI  
; APPLICANT: KOYAMA, MASARU  
; APPLICANT: MIYAKE, MASARUKI  
; APPLICANT: SEMOO, MASAHARU  
; TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND  
; TITLE OF INVENTION: PRODUCTION THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN  
; STREET: 130 WATER STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/254,493  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/865552  
; FILING DATE: 09-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 079996-1991  
; FILING DATE: 12-APR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 085396-1991  
; FILING DATE: 17-APR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 022321-1992  
; FILING DATE: 07-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RESNICK, DAVID S.  
; REGISTRATION NUMBER: 34235  
; REFERENCE/DOCKET NUMBER: 41777  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 523-6440  
; TELEFAX: (617) 523-6440  
; TELEX: 200291 STRE UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 687 base pairs  
; TYPE: nucleic acid

; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to genomic RNA  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: HUMAN  
; CELL TYPE: BREAST CARCINOMA  
; CELL LINE: ZR-75-1  
; US-08-254-493-2  
  
Alignment Scores:  
Pred. No.: 5,58e-23 Length: 687  
Score: 257.50 Matches: 64  
Percent Similarity: 43.72% Conservative: 44  
Best Local Similarity: 25.91% Mismatches: 94  
Query Match: 17.77% Indels: 45  
gaps: 6  
  
US-09-905-674-2 (1-270) x US-08-254-493-2 (1-687)  
QY 16 LysTYrIleuLeuPheSerTYrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGly 35  
| | | | |  
Db 31 AAATACCTGCTGTTGGATTTAACTTCACTCTTGCGCTGGCGGATTGCTGCTCTGCC 90  
QY 36 ValGIleuTrpAlaTrpSerGlu-----LysGIYValLeuSerAspLeuThrLys 52  
| | | | |  
Db 91 ATGGACATATGGCTCCGATTCGACTCTCAGACCAAGACATCTTCGAGCAAGAACTAAT 150  
QY 53 ValThrArgMetHisGlyIleAspProValIleuValIleuMetValGIYValAlaMet 72  
| | | | |  
Db 151 AATAATAATTCACCTTCTACACAGAGTCTATTTCTGATCGGAGCGCGCCCTCATG 210  
QY 73 PheThrIleuGlyPheAlaGlyCysValGIYValAlaLeuArgGluAsnIleCysLeuLeuAsn 92  
| | | | |  
Db 211 ATGCTGGTGGCTTCTCTGCTGCTGCTGCGGCTGTCGAGGAGTCCCGATCATCTGGA 270  
QY 93 PhePheCysGIYThrIleValIleuIlePhePheLeuGlnIleuAlaValAlaValIleuAla 112  
| | | | |  
Db 271 CTGTCCTTCGGCTTCCTCTGCTGCTGCTGATATTCGCCATTGAATGAGCTGCGCATCTGGA 330  
QY 113 PheIleuPheGlnAspTrpValAlaArgAspArgPheArgGluPhePheGluSerAsnIleLys 132  
| | | | |  
Db 331 TATTCGCCAAGAGAGAGTATTAGGAGAGTCCAGAGGTTTTCACAGGACACCTACAC 390  
QY 133 SerTYrArg---AspAspIleAspLeuGlnAsnIleuAspSerLeuGlnIySAlaAsn 151  
| | | | |  
Db 391 AAGCTGAACCAAGAGATGAGCCCAAGCGGGAACGCTGAAGACCATCCATGCTGTG 450  
QY 152 GlnCysCysGIYAlaTYrGlyProGluAspTrpAspLeuAsnValTYrPheAsnCysSer 171  
| | | | |  
Db 451 AACCTGCTGTTGGTGGCTGGGGGGGTGGA----- 480  
QY 172 GlyAlaSerTYrSerArgGlnIyScysGIYValProPheSerCysCysValProAspPro 191  
| | | | |  
Db 481 -----CAGTTTATCTCAGACATCTGC-----CCC 504  
QY 192 AlaGlnIySValIleAsnThrGlnCysGIYTYrAspValArgIleGlnLeuIySserLys 211  
| | | | |  
Db 505 AAGAAAGGACGTACTCGAAGC----- 525  
QY 212 TrpAspGluSerIlePheThr---LysGIYCysIleGlnAlaLeuGluSerTrpLeuPro 230  
| | | | |  
Db 526 -----TTACCGGTGAAGCTGCTCTGATGCCATCAAGAGAGTCTTCGAC 570  
QY 231 ArgAsnIleTYrIleValAlaGlyValPheIleAlaIleSerLeuGlnIlePheGly 250  
| | | | |  
Db 571 AATAAATTCACATCATCGGCGAGTGGCATCGCATTTGCCGTGATGATATTTGGC 630  
QY 251 IlePheLeuAlaArgThrLeu 257  
| | | | |  
Db 631 ATGATCTTCAGATGATGATCTTG 651

## RESULT 8

US-08-408-222B-2

Sequence 2, Application US/08408222B

Patent No. 5776727

GENERAL INFORMATION:

APPLICANT: Ikeyama, Shuichi

APPLICANT: Koyama, Masaru

APPLICANT: Miyake, Masayuki

APPLICANT: Senoo, Masaharu

TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dike, Bronstein, Roberts &amp; Cushman

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/408,222B

FILING DATE: 22-MAR-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/254,493

FILING DATE: 06-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP-079996-1991

FILING DATE: 12-APR-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP-085396-1991

FILING DATE: 14-APR-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP-022321-1992

FILING DATE: 07-FEB-1994

ATTORNEY/AGENT INFORMATION:

NAME: Resnick, David S.

REGISTRATION NUMBER: 34,235

REFERENCE/DOCKET NUMBER: 41777-DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 687 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to genomic RNA

HYPOTHEICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: human

CELL TYPE: breast carcinoma

CELL LINE: ZR-75-1

US-08-408-222B-2

## Alignment Scores:

Pred. No.:	5 58e-23	length:	687
Score:	257,50	Matches:	64
Percent Similarity:	43.72%	Conservative:	44
Best Local Similarity:	25.91%	Mismatches:	94
Query Match:	17.77%	Indels:	45
DB:	1	Gaps:	6

US-09-905-674-2 (1-270) x US-08-408-222B-2 (1-687)

QY 16 LysTyrLeuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValIlePheLeuGly 35

```
Db 31 AAATACCTGCTTCGAGATTAACTTCATCTTCGTGCGTCCGGGATGCTGCTGCCC 90
QY 36 ValGlyLeuTrpAlaTrpSerGlu-----LysGlyValLeuSerAspLeuThrLys 52
Db 91 ATGGACTATGGCTCCGATTCGATTCGACGACGACGATCTTCGACGACAAACTAAT 150
QY 53 ValThrArgMetHisGlyIleAspProValValLeuValLeuMetValGlyValIleMet 72
Db 151 AATAATATATTCACCTTCCTACACGAGGTCTATTTGTGATCGGAGCGCGCCCTCATG 210
QY 73 PheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeuLeuAsn 92
Db 211 ATGCTGGTGGGCTTCCTGCTGCTGCTGCGGCGCTGCGAGGATCCAGTCATGCTGGA 270
QY 93 PhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaValLeuAla 112
Db 271 CTGTTCTTGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330
QY 113 PheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGlnSerAsnIleLys 132
Db 331 TATTCACACAGATGAGTGTATTAAGAGAGTCCAGAGGATTTTCAAGACACCTACAC 390
QY 133 SerTyrArg---AspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAlaAsn 151
Db 391 AAGCTGAACCAACGATGACGCCGAGCGGAAACGCTGAAGACCATCCTATGCTTG 450
QY 152 GlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCysSer 171
Db 451 AACTGCTGTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 172 GlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCysValProAspPro 191
Db 481 -----CAGTTTATCTCAGACATCTGC-----CCC 504
QY 192 AlaGlnLysValAlaAsnThrGlnCysGlyTyrAspValArgIleGlnLeuLysSerLys 211
Db 505 AAGAAGCAGCTACGAAAC----- 525
QY 212 TrpAspGluSerIlePheThr---LysGlyCysIleGlnAlaLeuGluSerTrpLeuPro 230
Db 526 -----TTCCGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 570
QY 231 ArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIlePheGly 250
Db 571 AATAATTCACATCATCGGCGCATGCGCATGCGCATGCGCATGCGCATGATATTGCG 630
QY 251 IlePheLeuAlaArgThrLeu 257
Db 631 ATGATCTTCAGTATGATCTTG 651
```

## RESULT 9

US-08-408-222B-3

Sequence 3, Application US/08408222B

Patent No. 5776727

GENERAL INFORMATION:

APPLICANT: Ikeyama, Shuichi

APPLICANT: Koyama, Masaru

APPLICANT: Miyake, Masayuki

APPLICANT: Senoo, Masaharu

TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dike, Bronstein, Roberts &amp; Cushman

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

```
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,222B
FILING DATE: 22-MAR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,493
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-079996-1991
FILING DATE: 12-APR-1991
APPLICATION NUMBER: JP-085396-1991
FILING DATE: 14-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-022321-1992
FILING DATE: 07-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41777-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: breast carcinoma
CELL LINE: ZR-75-1
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 112...795
OTHER INFORMATION: E Mat peptide
US-08-408-222B-3

Alignment Scores:
Pred. No.: 1,18e-22 Length: 1120
Score: 257.50 Matches: 64
Percent Similarity: 43.72% Conserves: 64
Best Local Similarity: 25.91% Mismatches: 94
Query Match: 17.77% Indels: 45
DB: 1 Gaps: 6

US-09-905-674-2 (1-270) x US-08-408-222B-3 (1-1120)
QY 16 LysTyrLeuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGly 35
DB 142 AATATACCTGCTTGGATTAACTTCACTTCCTGCGGAGATCGCTCCCTCC 201
QY 36 ValGlyLeuTrpAlaTrpSerGlu-----LysGlyValLeuSerAspLeuThrIys 52
DB 202 ATTGACATATGCGTCGACATCTGACACAGACAGACATCTTGCAGCAAGAACTAAT 261
QY 53 ValThrArgMetHisGlyIleAspProValValLeuValLeuMetValGlyValValMet 72
DB 262 AATAATAATCCAGCTTACACAGAGATATATTCGATCGGAGCCGCCCTCATG 321
QY 73 PheThrLeuGlyPheAlaGlyCysValAlaLeuAlaGluAsnIleCysLeuLeuAsn 92
DB 322 ATGCTGGGGCTTCTGCGGCTGCGGCGGCTGTCAGAGGAGCCAGTCGATCGTGGGA 381
QY 93 PhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaValLeuAla 112
DB 382 CTGTTCTTCGCGCTTCCTTGTGTATATTCGCCCATTAAGTAAGTGGGCGCATCTGGGGA 441
```

```
QY 113 PheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGluSerAsnIleLys 132
DB 442 TATTCGCCACAGATGAGGTATTAAGGAAGTCCAGGAGTTTACAGACACCTACAAAC 501
QY 133 SerTyrArg-----AspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAlaAsn 151
DB 502 AAGCTGAACACCAAGAGATGAGCCCGGAGGAAACGCTGAAGACCATCCTATGCGCTG 561
QY 132 GlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCysSer 171
DB 562 AACTGCTGTGTTGGTGGGCGGCGTGGAA----- 591
QY 172 GlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCysValProAspPro 191
DB 532 -----CAGTTTATCTCAGACATCTGC-----CCC 615
QY 192 AlaGlnLysValAlaAsnThrGlnCysGlyTyrAspValArgIleGlnLeuLysSerLys 211
DB 616 AAGAGAGACGATCTGGAAC----- 636
QY 212 TrpAspGluSerIlePheThr---LysGlyCysIleGlnAlaLeuGluSerTrpLeuPro 230
DB 637 -----TTCACCGTGAAGTCTGCTCTGATGCCATCAAGAGAGCTCTGCAC 681
QY 231 ArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIlePheGly 250
DB 682 AATAAATTCACATCATCGCGGCGGAGTGGCGCATTCGCGTGCATGATGATATTTGGC 741
QY 251 IlePheLeuAlaArgThrLeu 257
DB 742 ATGATCTCTCAGTATGATCTTG 762

RESULT 10
US-08-430-225A-19
Sequence 19, Application US/08430225A
Patent No. 6204000
GENERAL INFORMATION:
APPLICANT: Dong, Jin-Tang; Barrett,
TITLE OF INVENTION: DIAGNOSTIC METHODS AND
TITLE OF INVENTION: GENE THERAPY USING REAGENTS DERIVED FROM THE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,225A
FILING DATE: 28-APR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4172
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
```

```

; TOPOLOGY: linear
US-08-430-225A-19

Alignment Scores:
Pred. No.: 2,086-22      Length: 1624
Score: 257.50           Matches: 67
Percent Similarity: 46.54% Conservative: 54
Best Local Similarity: 25.77% Mismatches: 116
Query Match: 17.77%      Indels: 23
DB: 4                    Gaps: 7

US-09-905-674-2 (1-270) x US-08-430-225A-19 (1-1624)

QY 16 LysTyrLeuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGly 35
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 193 AATATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 252

QY 36 ValGlyLeuTrpAlaTrpSerGluLys-----GlyValLeuSerAspLeuThr 51
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 253 TTCGGGGTGTGGATCTGCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312

QY 52 LysValThrArgMetHisGlyIleAspProValIleLeuValIleuMetValGlyValVal 71
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 313 AACTGCTTAGGATG-----GGGGCTATGCTTCTCATCGCGCTGGGGGCACTG 360

QY 72 MetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeuLeu 91
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 ACTATGCTATGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420

QY 92 AsnPhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaValLeu 111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 GGGCTGACTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480

QY 112 AlaPheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGluSerAsnIle 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 TTCTACTTCAACATGGGCAAGTGAAGCAAGAGATGGCGGCATCGTACTGAGCTCACT 540

QY 132 LysSerTyr-----ArgAspAspIleAspLeuGlnAsnIleAspSerLeuGln 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 541 CGAGACTACACAGACAGATCGCGAGAC---AGCTGACAGAGATGCTGGGACTACGTGAG 597

QY 149 LysAlaAsnGlnCysCysGlyAlaTrpGlyProGluAspTrpAspLeuAsnValTrpPhe 168
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 598 GCTCAGGTGAAGTGTGGGCTGCGCTCAGCTTCTACAACTGACAGACAGCGTGAGCTC 657

QY 169 AsnCysSerGlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCysCysVal 188
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 658 -----ATGATCGCCCTGAGTCACTCACTCCCTGTTCTGCGCAAGTC 699

QY 189 ProAspProAlaGlnLysValAlaAsnThrGlnCysGlyTyr-----AspVal 204
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 700 AAGGGGGAAGAGACACAGCCTTCTGTAGAGAGGCTTCTGAGAGCCGCCCGGCAAC 759

QY 205 ArgIleGlnLeu---LysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGln 223
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 760 AAGACCCGAGAGTGGCACACCTGAGGAGCTGCTGTACCAAGAGGCTGATGAG 819

QY 224 AlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValAlaIleAlaIle 243
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 820 AAGGTGCAAGGCTGCTGTCAGAGAACCTGGGCACTATCTCTGGGCTGGGCTGTG 879

QY 244 SerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGluAla 263
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 880 GCCATCATGACACTCTGCGGATGTCTGTCCATCTGCTGTGCGGCGAGTCCATTC 939

RESULT 11
; Sequence 2, Application US/08855140
; Patent No. 5854022
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.

```

```

; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: NEW TRANSMEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/855,140
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0296 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ. ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MYCONOT01
; CLONE: 779308
; US-08-855-140-2

Alignment Scores:
Pred. No.: 4,046-22      Length: 977
Score: 252.50           Matches: 61
Percent Similarity: 46.70% Conservative: 45
Best Local Similarity: 26.87% Mismatches: 90
Query Match: 17.43%      Indels: 31
DB: 2                    Gaps: 7

US-09-905-674-2 (1-270) x US-08-855-140-2 (1-977)

QY 13 CysTrpTyrLysTyrLeuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValAl 32
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 TGCCTTGAAGTGAACAGATGATCTCTCTCAATTTGATATTCCTGCTGTGG 199

QY 33 PheLeuGlyValGlyLeuTrpAlaTrpSerGluLysGlyValLeuSerAspLeuThrLys 52
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 CTGCTGGAGTGGGCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 253

QY 53 ValThrArgMetHisGlyIleAspProValIleLeuValIleuMetValGlyValAlaMet 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 254 TCCCCAGCTTCCCTGTTGTGTGTCGACGCCACCTGGTATCGCCAMPAGGACCACTTGTG 313

QY 73 PheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeuLeuAsn 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 314 ATGCTGACGGGCTTCTCGGCTCTGCGGGCCATCAGAGAAAAAAGAGGCTCTCTC 373

QY 93 PhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaValAla 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 374 AGCTTTTCAATGCTGCTGTGTGCTGACCCCTCTGACAGAGTGTATCTTACCACTCTTC 433

QY 113 PheLeuPheGlnAspTrpValAlaArgAspArgPheArgGluPhePheGluSerAsnIleLys 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 434 TTGTCTACATGACAGGTGAACGAGAACGCCAGACGACCTGAGAGAGCCCTGCTG 493  
 QY 133 SerTyr-----ArgaspAspIleaspLeuGlnaspLeuIleaspSerLeuGlnysala 150  
 Db 494 CTGTACCACACCGAACAACGCGGGCTGTAACAGCCTGGAACATCTCCAGAGGTGAG 553  
 QY 151 AsnGlnCysCysGlyAlaTyrGlyProGluAspTyrAspLeuAsnValTyrPheAsnCys 170  
 Db 554 ATCGCATGCTGTGCTGCTACTACACAGACTG-----TACCAGATG 598  
 QY 171 SerGlyAlaSerTyrSerArgGlnCysGlyValProPheSerCysCysValProAsp 190  
 Db 599 CTGGGGAGAAC-----ACGGTTCGCCAGCCGCTGTCATGAG----- 637  
 QY 191 ProAlaGlnIlyValAlaAsnThrGln---CysGlyTyrAspValArgIleGlnLeuLys 209  
 Db 638 -----AACTCCAGGGCTGGGGCGCAACGGCCAGCCCTTTG--- 676  
 QY 210 SerLysTyrAspGlnSerIlePheThrLysGlyCysIleGlnAlaLeuGlnSerTyrLeu 229  
 Db 677 -----TGGAGACGGGCTGCMATKAAAGGTGAAGATGTGGTTC 715  
 QY 230 ProArgAsnIleTyrIleVal 236  
 Db 716 GATGACAAATAGCAGCTGCTT 736  
 RESULT 12  
 US-08-254-493-3  
 ; Sequence 3, Application US/08254493  
 ; Patent No. 5439886  
 ; GENERAL INFORMATION:  
 ; APPLICANT: IKEYAMA, SHUICHI  
 ; APPLICANT: KOYAMA, MASARU  
 ; APPLICANT: MIYAKE, MASAYUKI  
 ; APPLICANT: SENO, MASAHARU  
 ; TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN  
 ; STREET: 130 WATER STREET  
 ; CITY: BOSTON  
 ; STATE: MASSACHUSETTS  
 ; COUNTRY: US  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/254,493  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/865552  
 ; FILING DATE: 09-APR-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 079996-1991  
 ; FILING DATE: 12-APR-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 085396-1991  
 ; FILING DATE: 17-APR-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 022321-1992  
 ; FILING DATE: 07-FEB-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: RESNICK, DAVID S.  
 ; REGISTRATION NUMBER: 34235  
 ; TELECOMMUNICATION INFORMATION: 41777  
 ; TELEPHONE: (617) 523-3400  
 ; TELEFAX: (617) 523-6440

; TELEX: 200291 STRE UR  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1120 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA to genomic RNA  
 ; HYPOTHEICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: HUMAN  
 ; CELL TYPE: BREAST CARCINOMA  
 ; CELL LINE: ZR-75-1  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 112..795  
 ; NAME/KEY: mat \_peptide  
 ; LOCATION: 115..795  
 ; US-08-254-493-3  
 Alignment Scores:  
 Pred. No.: 4.98e-22 Length: 1120  
 Score: 252.50 Matches: 63  
 Percent Similarity: 43.32% Conservative: 44  
 Best Local Similarity: 25.51% Mismatches: 95  
 Query Match: 17.43% Indels: 45  
 Db: 1 Gaps: 6  
 US-09-905-674-2 (1-270) x US-08-254-493-3 (1-1120)  
 QY 16 LysTyrIleLeuPheSerTyrAsnIleIlePheThrPheuAlaGlyValAlaPheLeuGly 35  
 Db 142 AAATACCTGCTGTTGGATTAACCTTCTTGCGCTGCCGAGATTCGCTGCTTCC 201  
 QY 36 ValGlyLeuThrPalatyrPserGlu-----LysGlyValLeuSerAspLeuThrLys 52  
 Db 202 ATGGACTATGAGCTCCGATTCGACTTCAGACCAAGACATCTTCGACCAAGAACTAAT 261  
 QY 53 ValThrArgMetHisGlyIleAspProValValLeuValLeuMetValGlyValValMet 72  
 Db 262 AATAATAATTCACAGCTTCTACACAGAGACTCTATATCTGATCGGAGCCGCCCTCATG 321  
 QY 73 PheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGlnAsnIleCysLeuLeuAsn 92  
 Db 322 ATGCTGTGGGCTTCTGTGGCTGCTGCGGCTGTGAGAGATCCAGTGTGATCTGGA 381  
 QY 93 PhePheCysGlyThrIleValIleLeuIlePhePheLeuGlnLeuAlaValAlaValLeuAla 112  
 Db 382 CTGTTCTTGGCTTCTGCTGCTGATATTCGCCATTGAAATAGCTGCGGCCATCTGGGGA 441  
 QY 113 PheLeuPheGlnAspTyrPvalArgAspArgPheArgGlnPhePheGlnSerAsnIleLys 132  
 Db 442 TATTCACCAAGAGACTGATTATTAAGAGAGTCCAGAGATTTCACAGACACCTACAC 501  
 QY 133 SerTyrArg---AspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAlaAsn 151  
 Db 502 AAGCTGAAACCAAGAGATGAGCCCGAGGGGAAACGCTGAAGACCATCTACATGCTTG 561  
 QY 152 GlnCysCysGlyAlaTyrGlyProGluAspTyrAspLeuAsnValTyrPheAsnCysSer 171  
 Db 562 AACTGCTGTGGTTGGCTGGGGCGGTGGA----- 591  
 QY 172 GlyAlaSerTyrSerArgGlnLysCysGlyValProPheSerCysCysValProAspPro 191  
 Db 592 -----CAGTTTACTCGACATCTGC-----CCC 615  
 QY 192 AlaGlnLysValAlaAsnThrGlnCysGlyTyrAspValArgIleGlnLeuLysSerLys 211  
 Db 616 AAGAGGACGTACTCGAAGAC----- 636  
 QY 212 TyrAspGlnSerIlePheThr---LysGlyCysIleGlnAlaLeuGlnSerTyrPhePro 230



Db 637 -----TTCCACCGTGAAGTCTGTCTGATGTCATCAAGAGGTCTTCGAC 681  
QY 231 ArgasnileTyrlleValAlaGlyValAlaPheIleAlaIleSerleuLeuGlnIlePheGly 250  
Db 682 AATRAATTCACATCATCGCCGCAAGTGGCATTCGCGTTCGATGATATTGTCG 741  
QY 251 IlePheLeuAlaArgThrLeu 257  
Db 742 ATGATCTTCAGTATGATCTTG 762  
RESULT 13  
US-08-808-148-2  
Sequence 2, Application US/08808148  
Patent No. 6020478  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer  
APPLICANT: Goli, Surya  
TITLE OF INVENTION: NOVEL HUMAN TUMOR-ASSOCIATED ANTIGEN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: IBM Compatible  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/808,148  
FILING DATE: Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0218 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 933 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSNOT26  
CLONE: 2187263  
US-08-808-148-2  
Alignment Scores:  
Pred. No.: 6,83e-18 Length: 933  
Score: 218.50 Matches: 72  
Percent Similarity: 44.09% Conservative: 51  
Best Local Similarity: 25.81% Mismatches: 112  
Query Match: 15.08% Indels: 45  
DB: 3 Gaps: 8  
US-09-905-674-2 (1-270) x US-08-808-148-2 (1-933)  
QY 7 SerAsnAlaLysValSerCys-----TrpTyrIysTyrLeuLeuPheSerTyrAsnIle 24  
Db 116 ACAGAGCACCACCATGATGCTTCAGCTTCAATTAGACCATGATGATGCTCTTCATTTG 175

QY 25 IlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAlaTrpSerGlnLys 44  
Db 176 CTCATCTTTCTGTGTGGTGCACCCCTGTTGGCAGGGCATCTGGGTGC-AATCGATGG 234  
QY 45 GlyValLeuSerAsp-----LeuThrLysVal-ThrArgMetHisGlyIleAsp 61  
Db 235 GGCATCTCTTGAAGATCTTCGGGCCACTGTCGTCAGTGCATGCAATGCTTTCACACT 294  
QY 61 oValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCysVa 81  
Db 295 GGGCTACTTCTCATCGCAGCCGCGGTGTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTA 354  
QY 81 lGlyValLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuI 101  
Db 355 TGTCTCTAAGACTGAGACAAAGTGGCCCTGTCGACGTTCTTCATCTTCCTCATCTCAT 414  
QY 101 ePhePheLeuGlnLeuAlaValAlaValAlaPheLeuPheGlnAspTrpValArgAs 121  
Db 415 CTTCATTCCTGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459  
QY 121 parPheArgGluPhePheGlnSerAsnIle-----LysSerTyrAr 135  
Db 460 CACAATGGCTGAGCACTTCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519  
QY 135 gAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGly 155  
Db 520 TTCCAGGAAGACTTTCACCTCAAGTGTGGACACACCACTGAAGAGCTCAAGTCTGTGG 579  
QY 155 yAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTy 175  
Db 580 CTTCACCAACTATACGATTTTGAAGACTCACCTACTTC----- 619  
QY 175 rSerArgLysCysGlyValProPheSerCysCys-----ValProAspProAl 192  
Db 620 ----AAAGAGAACAGTGCCTTCCCATCTCTGTCATGATGACCAACCTCAACACAGC 675  
QY 192 aGlnLysValAlaAsnThrGlnCysGlyTyrAspValArgIleGlnLeuLysSerLysTr 212  
Db 676 CAATGAACCTGCACCAACGAAAGGCTCAGACACCAAAAGCTAGAG----- 721  
QY 212 pAspGlnSerIlePheThrLysGlyCysIleGlnAlaLeuGlnSerTrpLeuProArgAs 232  
Db 722 -----GGTTCCTCAATCAGCTTTTGTATGATCATGCACTAA 759  
QY 232 nIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIlePheGlyIlePh 252  
Db 760 TGCAGTCACCGTGGGTGGTGGCAGCTGGAATTGGGGCTCGACCTGCTGCCATGAT 819  
QY 252 eleuAlaArgThrLeuIleSerAspIleGluAlaValLysAlaGlyHisHisPhe 270  
Db 820 TGTGTCATGTATCTGACTGCATCATCAATAAGTC-----CACTTC 862  
RESULT 14  
US-09-020-956-111  
Sequence 111, Application US/09020956  
Patent No. 6261562  
GENERAL INFORMATION:  
APPLICANT: Xu, Jianshun  
APPLICANT: Dillin, David C.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS  
NUMBER OF SEQUENCES: 178  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/020,956  
 FILING DATE: 09-FEB-1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MAKI, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.427C2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 111:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1289 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 US-09-020-956-111

Alignment Scores:  
 Pred. No.: 1,12e-17 Length: 1289  
 Score: 218.50 Matches: 72  
 Percent Similarity: 44.09% Conservative: 51  
 Best Local Similarity: 25.81% Mismatches: 112  
 Query Match: 15.08% Indels: 45  
 DB: 4 Gaps: 8

US-09-905-674-2 (1-270) x US-09-020-956-111 (1-1289)

QY 7 SeranAlaIysValSerCys-----TrrpYlYsTyrLeuLeuPheSerTyrAsnIle 24  
 Db 111 ACAGAGCCACCATGAGGCTTCACTTAAAGACATGATGCTTCAATTG 170  
 QY 25 IlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAlaTrpSerGluLys 44  
 Db 171 CTAATCTTCTGTGTGTGACACCTGTGTGGCAGTGGGCTGTC-AATGATGG 229  
 QY 45 GlyValLeuSerAsp-----LeuThrIysVal-ThrArgMetHisGlyIleAsp 61  
 Db 230 GGCATCTTCTGAGAGATCTTGGGCACTGTCTCCAGTGCACATGCTTGCACAGT 289  
 QY 61 oValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCysVa 81  
 Db 290 GGGCTACTCTCTATCGACGCGCGGTGTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 349  
 QY 81 lGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuI 101  
 Db 350 TGTGTCTAGACTGAGAGAGAGTGTGCTGTGACGTTCTTCTTCATCTCTCTCTCAT 409  
 QY 101 ePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAsnPTPrValArgS 121  
 Db 410 CTTTCATCTGTGAGGTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 454  
 QY 121 parPheArgGluPhePheGluSerAsnIle-----LysSerTyrAr 135  
 Db 455 CACAATGGCTGAGCATCTTCTGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 514  
 QY 135 gasPasPleAspLeuGlnAsnLeuIleAspSerLeuGlnIysAlaAsnGlnCysGly 155  
 Db 515 TTCCTCAGGAAGCTTCACTCAAGTGTGAGAACACACACATGAAAGGGCTCAAGTGTGTG 574  
 QY 155 yAlaTrpGlyProGluAspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyr 175  
 Db 575 CTTTCACCACTATACGATTTTGAAGACTCACCTACTTTC----- 614  
 QY 175 rSerArgGluLysCysGlyValAlaProPheSerCysCys-----ValProAspProAl 192  
 Db 615 ----AAAGAGACAGTGTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 670  
 QY 192 agInLysValIAsnThrGlnCysGlyTyrAspValAlaGlnLeuLysSerIysTr 212

Db 671 CATGAACCTGCACCAACCAAGGCTCACCAACCAAGGCTCACCAACCAAGGCTCACCAAC 716  
 QY 212 pasPgluSerIlePheThrIysGlyCysIleGlnAlaLeuGluSerTrpLeuProArgS 232  
 Db 717 -----GGTTGCTTCAATCAGCTTTTGTATGATCATCCGAACCTAA 754  
 QY 232 nIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIlePheGlyIlePh 252  
 Db 755 TGCAGTACCCGTGGT 814  
 QY 252 eLeuAlaArgThrLeuIleSerAspIleGluAlaValIysAlaGlyHisHisPhe 270  
 Db 815 TGTGTCAATATCTGTACTGCAATCTACAAATTAATC-----CACTTC 857

RESULT 15  
 US-09-030-607-111  
 Sequence 111, Application US/09030607  
 Patent No. 6262245  
 GENERAL INFORMATION:  
 APPLICANT: Xu, Jiangchun  
 TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS  
 NUMBER OF SEQUENCES: 224  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: SEED AND BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 City: Seattle  
 STATE: WA  
 COUNTRY: USA  
 ZIP: 98104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/030,607  
 FILING DATE: 25-FEB-1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MAKI, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.427C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 111:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1289 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 US-09-030-607-111

Alignment Scores:  
 Pred. No.: 1,12e-17 Length: 1289  
 Score: 218.50 Matches: 72  
 Percent Similarity: 44.09% Conservative: 51  
 Best Local Similarity: 25.81% Mismatches: 112  
 Query Match: 15.08% Indels: 45  
 DB: 4 Gaps: 8

US-09-905-674-2 (1-270) x US-09-030-607-111 (1-1289)

QY 7 SeranAlaIysValSerCys-----TrrpYlYsTyrLeuLeuPheSerTyrAsnIle 24  
 Db 111 ACAGAGCCACCATGAGGCTTCACTTAAAGACATGATGCTTCAATTG 170  
 QY 25 IlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAlaTrpSerGluLys 44

Db	171	CTCATCTTCTGTGGTGGAGCCCTGTTGGCATGGGCAATCTGGGTCT-AATCATGG	229
QY	45	GLYValLeuSerASP-----LeuThrLysVal-ThrArgMetHisGLYLeuAsp	61
Db	230	GGCATCTCTTTTGGAAAGATCTTCGGGCCACTGCTCCACAGTGCACATGCAGTTTGTCAACGT	289
QY	61	oValValLeuValLeuMetValGLYValValMetPheThrLeuGlyPheHisGLYVal	81
Db	290	GGGCTACTTCTCTATCCGACGCCGCCGCTTGTGTGCTTTGCTCTTGTGTTCCTGGGCTGCTA	349
QY	81	IGLYValLeuArgGLYAsnLysLysLysLeuLeuAsnPhePheCysGLYThrIleValLeuI	101
Db	350	TGGTGTAAAGCTAGAGCAAGAGTGTGCCCTGTGACGTCTTCTCATCTCCTCCAT	409
QY	101	ePhePheLeuLeuLeuAlaValAlaValIleLeuAlaPheLeuPheGLYAsnIleValArg	121
Db	410	CTTATTTGCTGAGGTGTCAAGCTGTGTGGTGTGCGCTTGTGTATC-----AC	454
QY	121	PaArgPheArgLysPheGlyLeuSerAsnIle-----LysSerTyr	135
Db	455	CACATAGGCTGAGCATCTCCGACAGTTGTGTGATGTCCCTGCATCAAGAAACATATATGG	514
QY	135	GLSPAspIleAspLeuGLYAsnLeuIleAspSerLeuGLYAlaAsnGLYCysGL	155
Db	515	TTTCCAGGAACACTTCACTCAAGTGTGAACACCACTCAAGAAAGGCTCAAGTGTGTGG	574
QY	155	YAlaTyrGLYProLysPhePheAspLeuAsnValTyrPheAsnLysSerGLYAserTyr	175
Db	575	CTTACCAACATATACGATTTTGGAGCTACACCTACTTC-----	614
QY	175	rSerArgGLYLysCysGLYValProPheSerCys-----ValProAspProAl	192
Db	615	-----AAAGAGAACAGTCTCTTCCCATCTGTGTGAATGACACAGTACCAACACAGC	670
QY	192	AGLysValValAsnThrGlyGlyTyrTrpAspValArgIleGlnLeuLysSerLysTr	212
Db	671	CAATGAAGAACTGCACCAAGCAAAAGGCTCCAGCAAAAGGTAGAG-----	716
QY	212	PaSPGLuSerIlePheThrLysGLYCysIleGlnAlaLeuGLYSerTrpLeuProArg	232
Db	717	-----GGTTCCTCAATCACTTGTATGACATCCGAACATA	754
QY	232	nIleTyrIleValAlaGLYValPheIleAlaIleSerLeuLeuGlnIlePheGLYLeuP	252
Db	755	TGCAGTCCACGCTGGGTGGTGGCAGCTGAATTTGGGGCTGCAGCTGCATCATGAT	814
QY	252	eLeuAlaArgThrLeuLeuSerAspIleGlyAlaValLysAlaGLYHisHisPhe	270
Db	815	TGTGTCCATGATATCTGTACTGCATATCAATATAAGTC-----CACTGT	857

Search completed: December 19, 2002, 07:46:59  
Job time : 56 secs













|||||  
Db 965 CAGTGAAGCCGCGCATCATCTCTAGAGAGCAGAGCTTAGAGGAGCCGAGCTGAGCCACGC 1024  
OY 914 TGGGAGCCAGAGACCTTTCTCTGACATCAGCCCTACGTCAGAGGAGAGAGAGCCGAGAC 973  
Db 1025 TGGGAGCCAGAGACCTTTCTCTGACATCAGCCCTACGTCAGAGGAGAGAGAGCCGAGAC 1084  
OY 974 CCCCAGAGCAGTGGCCCCATCTTAAGCATCAGCGTAGCGTCACTCTCTGTTTCTGCTTG 1033  
Db 1085 CCCAGAGCCAGTGGCCCCATCTTAAGCATCAGCGTAGCGTCACTCTCTGTTTCTGCTTG 1144  
OY 1034 CTGGTCTTAAGACCAAGGCTCCCTTGTACTGCCCCAACTTGTACTGATCCCTC 1093  
Db 1145 CTGGTCTTAAGACCAAGGCTCCCTTGTACTGCCCCAACTTGTACTGATCCCTC 1204  
OY 1094 TGGAGTCTACCCAGAGACAGAGATGTGCTTATGTGGAGTGGTGACTCGAAGAC 1153  
Db 1205 TGGAGTCTACCCAGAGACAGAGATGTGCTTATGTGGAGTGGTGACTCGAAGAC 1264  
OY 1154 GAGAGGGCTCTGTGCTGCCAGAGAGGCTTGACTCAGACCCCTGACGCTCAAGCATGT 1213  
Db 1265 GAGAGGGCTCTGTGCTGCCAGAGAGGCTTGACTCAGACCCCTGACGCTCAAGCATGT 1324  
OY 1214 CTGCAGAGACCCCTGCTGCCCTCTCCATGTGCGTCCAGACATCTGTTGGGTATCCA 1273  
Db 1325 CTGCAGAGACCCCTGCTGCCCTCTCCATGTGCGTCCAGACATCTGTTGGGTATCCA 1383  
OY 1274 CATCTGTGGTNGGCCGTGGGTAGAGGGACCCACAGGGCTGAGACGGGATCTCTCTCCA 1333  
Db 1384 CATCTGTGGTNGGCCGTGGGTAGAGGGACCCACAGGGCTGAGACGGGATCTCTCTCCA 1443  
OY 1334 TCAGCAAAAGCAGCATGGGGGCTTGGCTTAACGGGAGCG 1376  
Db 1444 TCAGCAAAAGCAGCATGGGGG--CCTGCCGTAAAGGGAGCG 1484  
  
RESULT 4  
AAFP0628  
ID AAFP0628 standard; cDNA: 2715 BP.  
XX  
AC AAFP0628;  
XX  
DT 04-MAY-2001 (first entry)  
XX  
DE Human TANGO 339 cDNA, SEQ ID NO:1.  
XX  
KW Secreted protein; transmembrane protein; TANGO; human; drug screening;  
KW activity modulator; expression modulator; cancer; immunological disorder;  
KM cytosolic; immunomodulatory; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200109162-A2.  
XX  
PD 08-FEB-2001.  
XX  
PF 31-JUL-2000; 2000WO-US20935.  
XX  
PR 30-JUL-1999; 99US-0365164.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Fraser CC, Sharp JD, Kirst SJ, Barnes TW, Wrighton N, Myers PS;  
PI Pan Y;  
XX  
DR WPI: 2001-138647/14.  
DR P-PSDB: AAB87034, AAB87035, AAB87036.  
XX  
PT Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful  
PT for the prevention, diagnosis and treatment of, e.g. cancers and immune  
PT disorders -  
XX  
PS Claim 2; Page 219-220; 332pp; English.

XX  
CC The invention relates to novel secreted/transmembrane proteins, and  
CC nucleic acids encoding them. The novel proteins are designated TANGO 339,  
CC TANGO 353, TANGO 358, TANGO 365, TANGO 366, TANGO 369, TANGO 383, TANGO  
CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a  
CC murine TANGO 393 is also included within the scope of the invention. The  
CC invention also encompasses fragments and variants of the proteins of the  
CC invention, and nucleic acids encoding them. The invention additionally  
CC relates to host cells comprising a nucleic acid of the invention; methods  
CC for the production of a protein of the invention; an antibody specific  
CC for a protein of the invention; methods for detecting a protein or  
CC nucleic acid of the invention; and methods of identifying agents which  
CC bind to or modulate the activity of a protein of the invention. The novel  
CC secreted proteins, nucleic acids encoding them, and antibodies against  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression or activity of the secreted  
CC proteins. The secreted proteins of the invention may also be used to  
CC identify modulators of expression or activity, which may be useful in  
CC the treatment of disorders associated with the proteins of the  
CC invention e.g., cancers and immunological disorders. The present  
CC sequence represents cDNA encoding human TANGO protein of the invention.  
XX  
SQ Sequence 2715 BP; 559 A; 760 C; 715 G; 681 T; 0 other;

Query Match 93.1%; Score 1292.2; DB 22; Length 2715;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1317; Conservative 2; Mismatches 6; Indels 3; Gaps 2;

OY 49 GATTCTGCTTCTCAGAAAGATGACATATTATAGATCTTAACGCCAAAGTACAGTCTGTG 108  
Db 192 GATCTCTGCTTCTCAGAAAGATGACATATTATAGATCTTAACGCCAAAGTACAGTCTGTG 251  
OY 109 TACAAGTACCTCCCTTTACACTCAACATATCTTCTGTTGGCTGAGATGTTCTCTT 168  
Db 252 TACAAGTACCTCCCTTTACACTCAACATATCTTCTGTTGGCTGAGATGTTCTCTT 311  
OY 169 GGAATCGGCGCTGGGCGATGAGCGAAAGGCTGCTGCCACCTCAACAAAGTGACC 228  
Db 312 GGAATCGGCGCTGGGCGATGAGCGAAAGGCTGCTGCCACCTCAACAAAGTGACC 371  
OY 229 CGGATGATGGAATCGACCCCTGTGCTGCTCTGATGAGTGGGCTGTGATGTTCCACC 288  
Db 372 CGGATGATGGAATCGACCCCTGTGCTGCTCTGATGAGTGGGCTGTGATGTTCCACC 431  
OY 289 CTGGGGTTCGCGCGCGCTGGGCGCTTCCGCGGAGAAATATCTGCTTGGCTCAACTTTTC 348  
Db 432 CTGGGGTTCGCGCGCGCTGGGCGCTTCCGCGGAGAAATATCTGCTTGGCTCAACTTTTC 491  
OY 349 TGTGGCACATCGTGCATCTTCTTCGAGACTGAGTGGCGCTGCTGGGCTTCTCTG 408  
Db 492 TGTGGCACATCGTGCATCTTCTTCGAGACTGAGTGGCGCTGCTGGGCTTCTCTG 551  
OY 409 TTCCAGAGACTGGGTGAGGAGCCGGTTCGGGAGTTCTTCAGAGCAACATCAAGTCTTAC 468  
Db 552 TTCCAGAGACTGGGTGAGGAGCCGGTTCGGGAGTTCTTCAGAGCAACATCAAGTCTTAC 611  
OY 469 CGGAGCATATGATCTGCAAAACCTCATGAGTCTCCCTTCAGAAAGCTTAACAGTCTCT 528  
Db 612 CGGAGCATATGATCTGCAAAACCTCATGAGTCTCCCTTCAGAAAGCTTAACAGTCTCT 671  
OY 529 GGGCATATGGCCCTGAAGACTGGACCTCAAGTCTACTTCAATTCAGAGCGGTGGCAGC 588  
Db 672 GGGCATATGGCCCTGAAGACTGGACCTCAAGTCTACTTCAATTCAGAGCGGTGGCAGC 731  
OY 589 TACAGCCGAGAGAAAGTGGGGTCCCTTCTCTGCTGCTGCGAGATCTTGCGCAAAA 648  
Db 732 TACAGCCGAGAGAAAGTGGGGTCCCTTCTCTGCTGCTGCGAGATCTTGCGCAAAA 791  
OY 649 GTTGTGAACACAGTGTGATATGATGAGGATTCAGTGAAGGCAAGTGGGTGAG 708  
Db 792 GTTGTGAACACAGTGTGATATGATGAGGATTCAGTGAAGGCAAGTGGGTGAG 851  
OY 709 TCCATCTTCAAGAAAGCTGCATCCAGGGCTGGAAGAGTGGCTCCCGCGGAACATTTAC 768

```
Db 852 TCGATCTTCAGAAAGCGTCGATCCAGCGCTGGAAAGCTGCTCCCGGAAACATTAC 911
Oy 769 ATTGTCGCTGGCGCTTTCATCCCATCTCGCTGTTGCAGATATTGGCATCTTCTGCA 828
Db 912 ATTGTCGCTGGCGCTTTCATCCCATCTCGCTGTTGCAGATATTGGCATCTTCTGCA 971
Oy 829 AGGACGCTGATCTCGACATCGAGGAGTGAAGCGCGCCCTTCATCTTCTGAGAGAG 888
Db 972 AGGACGCTGATCTCGACATCGAGGAGTGAAGCGCGCCCTTCATCTTCTGAGAGAG 1031
Oy 889 TTGAGGAGCCGAGCTGAGCCAGCGCTGGAGGCCAGAGCCCTTCTCTCCATCAGCCCTA 948
Db 1032 TTGAGGAGCCGAGCTGAGCCAGCGCTGGAGGCCAGAGCCCTTCTCTCCATCAGCCCTA 1091
Oy 949 GGTCCAGAGGAGAGAGAGAGCCAGACCCCAAGAGCCAGTGGCCCATCTTAAGATCAGCGT 1008
Db 1092 GGTCCAGAGGAGAGAGAGCCAGACCCCAAGAGCCAGTGGCCCATCTTAAGATCAGCGT 1151
Oy 1009 GACGTGACCTCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
Db 1132 GACGTGACCTCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
Oy 1069 GCCCAACTTGTGACTGATCCCTCTGAGTCTACCCAGAGACAGAGATGTCTTTAT 1128
Db 1212 GCCCAACTTGTGACTGATCCCTCTGAGTCTACCCAGAGACAGAGATGTCTTTAT 1271
Oy 1129 GTGGAGTGTGACTGATCCCAAGAGAGAGAGGCTCTGCTGCTGCTGCTGCTGCTGCT 1188
Db 1272 GTGGAGTGTGACTGATCCCAAGAGAGAGAGGCTCTGCTGCTGCTGCTGCTGCTGCT 1331
Oy 1189 CAGACCCCTGAGCTCAAGCATGTCTGTCAGAGACACCCCTGCTCCCTCTCCATTTGCGWT 1248
Db 1332 CAGACCCCTGAGCTCAAGCATGTCTGTCAGAGACACCCCTGCT - CCGCTCTCAGTGGAT 1390
Oy 1249 CCAGACATCTGCTTGGGTATCCACATCTGTGGGTGCGGTGAGAGAGAGCCACA 1308
Db 1391 CCAGACATCTGCTTGGGTATCCACATCTGTGGGTGCGGTGAGAGAGAGCCACA 1450
Oy 1309 GCGGTGAGAGAGGCGATCTCTCCATCAAGAGAGAGAGGCGGCGCTTGGCGTAAAC 1368
Db 1451 GCGGTGAGAGAGGCGATCTCTCCATCAAGAGAGAGAGGCGGCG - CTTGCCCGTAAAC 1508
Oy 1369 GGGAGGCG 1376
Db 1509 GGGAGGCG 1516

RESULT 5
AAS92016
ID AAS92016 standard; cDNA; 2623 BP.
XX
AC AAS92016;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #27820.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
```

```
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI, 2001-639362/73.
XX
DR P-PDB; ABG27829.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID NO 27820; 103bp; English.
XX
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 2623 BP; 515 A; 713 C; 719 G; 659 T; 17 other:
XX
XX
Query Match 91.9%; Score 1275.4; DB 23; Length 2623;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1345; Conservative 2; Mismatches 40; Indels 8; Gaps 4;
Oy 1 CTTCTCTGGGCGGAGCGGCGGCGGCGCTCCCGCGCGCGCGGCGGATCTGCTTCT 60
Db 1 CTTCTCTGGGCGGAGCGGCGGCGGCGGCGCTCCCGCGCGCGGCGGATCTGCTTCT 60
Oy 61 CAGAGATGCACTATTATGATATCTTAAGCCCAAGTACAGCTGTGTGATCAAGTACTTC 120
Db 61 CAGAGATGCGCTATGATATCTTAAGCCCAAGTACAGCTGTGTGATCAAGTACTTC 120
Oy 121 CTTTTCAGTACAACATCATCTTCTGTTGGCTGAGATTGTCTTCTTGGAGTGGGCTG 180
Db 121 CTTTTCAGTACAACATCATCTTCTGTTGGCTGAGATTGTCTTCTTGGAGTGGGCTG 180
Oy 181 TGGGCAATGAGCGAAGAGGCTGTGCTGTCCGACTTACCAACAAATGACCCGGATCGATGA 240
Db 181 TGGGCAATGAGCGAAGAGGCTGTGCTGTCCGACTTACCAACAAATGACCCGGATCGATGA 240
Oy 241 ATGCACCTGTGGTGGCTGCTGATGATGAGGCTGATGATGATGATGATGATGATGATGATG 300
Db 241 ATGCACCTGTGGTGGCTGCTGATGATGAGGCTGATGATGATGATGATGATGATGATGATG 300
Oy 301 GGTGCTGAGGAGGCTGCGGAGAAATATCTGCTTACCACTTTTCTGTGGACATTC 360
Db 301 GGTGCTGAGGAGGCTGCGGAGAAATATCTGCTTACCACTTTTCTGTGGACATTC 360
Oy 361 GTGCTCATCTTCTTCTGAGAGTGTGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 GTGCTCATCTTCTTCTGAGAGTGTGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Oy 421 GTGAGGAGCGGATTCGGGAGTCTTCGAGAGCAATCAAGTCTACCGGAGCATATC 480
Db 421 GTGAGGAGCGGATTCGGGAGTCTTCGAGAGCAATCAAGTCTACCGGAGCATATC 480
```

```
OY 481 GATCTGCAAAACCTCATGACTCCCTTCAGAAAGCTAACAGTGTGTGGCGCATATGCG 540
DB 481 GATCTGCAAAACCTCATGACTCCCTTCAGAAAGCTAACAGTGTGTGGCGCATATGCG 540
OY 541 CCTGAAGACTGGGACCTTAAGCTTACTTCAATTGACAGCGGTGCCAGTACAGCCGAGAG 600
DB 541 CCTGAAGACTGGGACCTTAAGCTTACTTCAATTGACAGCGGTGCCAGTACAGCCGAGAG 600
OY 601 AAGTCCGGGGTCCCTTCTGCTGGGTGGGTGCATGCTGGGCAAAAGTTGTGAACACA 660
DB 601 AAGTCCGGGGTCCCTTCTGCTGGGTGGGTGCATGCTGGGCAAAAGTTGTGAACACA 660
OY 661 CAGTGTGATATGATATGATGAGATTACAGTGAAGAGCAAGTGGATGATCCATCTTACG 720
DB 661 CAGTGTGATATGATATGATGAGATTACAGTGAAGAGCAAGTGGATGATCCATCTTACG 720
OY 721 AAAGGCTGCATTCAGAGCGCTGAAAAGCTGCTCCCGGGAAACATTTACATTTGTGGTGGC 780
DB 721 AAAGGCTGCATTCAGAGCGCTGAAAAGCTGCTCCCGGGAAACATTTACATTTGTGGTGGC 780
OY 781 GTCTTCATGCGCATCTGCTGCTGTCAGATATTTGGCANTCTTCCGCAAGAGACCTGATC 840
DB 781 GTCTTCATGCGCATCTGCTGCTGTCAGATATTTGGCANTCTTCCGCAAGAGACCTGATC 840
OY 841 TCAGACATCGAGGAGTGAAGCGCGGCATCTTCTGAGAGAGAGATTGAGGAGAGCGG 900
DB 841 TCAGACATCGAGGAGTGAAGCGCGGCATCTTCTGAGAGAGAGATTGAGGAGAGCGG 900
OY 901 AGCTGAGCCACGCTGGGAGGCGCAAGGCTTCTCTGCCATCAGCCCTACGTCCAGAGGGA 960
DB 901 AGCTGAGCCACGCTGGGAGGCGCAAGGCTTCTCTGCCATCAGCCCTACGTCCAGAGGGA 960
OY 961 GAGGAGCGGACACCCCGACAGCCAGTCCCATCTTAAGCATCAGCGT -GACGTGACTCTC 1019
DB 961 GAGGAGCGGACACCCCGACAGCCAGTCCCATCTTAAGCATCAGCGT -GACGTGACTCTC 1020
OY 1020 TCTGTTTC-TGCTTCTGCTGCTGCTGAGACCAAGGCTCCCTTGTACTCTCCCAAACTT 1078
DB 1020 TCTGTTTC-TGCTTCTGCTGCTGCTGAGACCAAGGCTCCCTTGTACTCTCCCAAACTT 1080
OY 1079 GTGACTGATCCCTCTGTGAGTCTACCCAGAGACAGAGATGTGCTTTATGTGGAGTGG 1138
DB 1081 GTGACTGATCCCTCTGTGAGTCTACCCAGAGACAGAGATGTGCTTTATGTGGAGTGG 1140
OY 1139 TGACTCTGAAGACAGAGAGGCTCTCTGTG-----CTGCCAGAGAGGCTTGACTCAGAC 1193
DB 1141 TGACTCTGAAGACAGAGAGGCTCTCTGTG-----CTGCCAGAGAGGCTTGACTCAGAC 1200
OY 1194 CCCCTGACAGTCAACATGCTGTGAGAGACCCCTGCTCCCTCTCCAYTGACGTCACAGA 1253
DB 1201 CCCCTGACAGTCAACATGCTGTGAGAGACCCCTG-8CCCCCTCTCCATGCTGACATCCAGA 1259
OY 1254 CATCTGCTTTGGGTGATCCACATCTGTGGTGGGTAGAGGAGCCACAGAGGCT 1313
DB 1260 CATCTGCTTTGGGTGATCCACATCTGTGGTGGGTAGAGGAGCCACAGAGGCT 1319
OY 1314 GGACAGGCGATCTCTCTCATCAACAAAGACAGCATGGGGGGCTTGGCCGTAACGGGAG 1373
DB 1320 GGACAGGCGATCTCTCTCATCAACAAAGACAGCATGGGGGGCTTGGCCGTAACGGGAG 1379
OY 1374 GCGNGACGTTGGGCC 1388
DB 1380 GCGNGACGTTGGGCC 1394
```

```
RESULT 6
AAZ65350
ID AAZ65350 standard; DNA; 1667 BP.
XX
AC AAZ65350;
XX
DT 23-MAR-2000 (first entry)
XX
```

```
DE Human secreted protein gene 10 fragment.
XX
KW Human: secreted protein; cancer; tumour; developmental abnormality;
KW foetal deficiency; blood disorder; immune system disorder; inflammation;
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KW therapy; chromosome 10; ds.
XX
OS Homo sapiens.
XX
PN W0958660-A1.
XX
PD 18-NOV-1999.
XX
PF 06-MAY-1999; 99MO-US09847.
XX
PR 12-MAY-1998; 98US-0085093.
PR 12-MAY-1998; 98US-0085094.
PR 12-MAY-1998; 98US-0085105.
PR 12-MAY-1998; 98US-0085180.
PR 18-MAY-1998; 98US-0085906.
PR 18-MAY-1998; 98US-0085920.
PR 18-MAY-1998; 98US-0085921.
PR 18-MAY-1998; 98US-0085922.
PR 18-MAY-1998; 98US-0085923.
PR 18-MAY-1998; 98US-0085924.
PR 18-MAY-1998; 98US-0085928.
PR 18-MAY-1998; 98US-0085925.
PR 18-MAY-1998; 98US-0085927.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
PI Lafleur DW, Endress GA, Ebner R;
XX
DR WPI: 2000-062296/05.
XX
P-PSDB: AAY76266.
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
XX
PS Disclosure: Page 427-428; 475pp: English.
XX
CC AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
CC This sequence was found to be present on human chromosome 10.
CC AAY76124 to AAY76223 represent the secreted proteins encoded by the 97
CC human genes. The genes and their corresponding secreted polypeptides are
CC useful for preventing, treating or ameliorating medical conditions,
CC e.g. by protein or gene therapy. Also pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new genes. Specific
CC uses are described for each of the 97 genes, based on which tissues they
CC are most highly expressed in, and include developing products for the
CC diagnosis or treatment of cancer, tumours, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
CC disorders, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The sequences shown in AAY76224 to AAY76424 represent fragments of the
CC secreted proteins.
XX
SQ Sequence 1667 BP; 323 A; 469 C; 488 G; 387 T; 0 other;
```

```
Query Match 85.7%; Score 1189.6; DB 21; Length 1667;
Best Local Similarity 95.8%; Pred. No. 36-303;
Matches 1309; Conservative 1; Mismatches 41; Indels 16; Gaps 8;
```

OY	18	GGGCGGGGGCCGGTGCACCCCCCGGGGGGGAGATTGCTTCACAAAGATGACATATA	77
Db	1	GGCCGCGCCCGGTGCGCGCGCGCGC-GATTCTGCTTCACAAGAATGACTATTTA	59
OY	78	TAGATACTCTAACGCCCAAAAGTCAGCTCTGGTACAAAGTACTCTTTTGAGTACAACT	137
Db	60	TAGATACTCTAACGCCCAAAAGTCACTCTGGTACAAAGTACTCTTTTGAGTACAACT	119
OY	138	CATCTTCTGGTTGGCTGGAGTGTCTTCTTGGAATCGGGCTGTGGCATGAGCGAANA	197
Db	130	CATCTTCTGGATTGGCTGGAGTGTCTTCTTGGAAGTGGGCTGTGGCATGAGCGAANA	179
OY	198	GGGTGCGTGTCCGACCACCAAAGTAGCCCGGAATGACGAATGAGACCTGGTGGCT	257
Db	180	GGGTGTCTGTCCGACCTCACCAAAGTAGCCCGGAATGACGAATGAGACCTGTGGTGT	239
OY	258	GGTCTGATGGTGGGCGGTGGTGGATGTTCACCTTGGGGTGTGCCGGCTGGGGGCTCT	317
Db	240	GGTCTGATGGTGGGCGGTGGTGGATGTTCACCTTGGGGTGTGCCGGCTGGGGGCTCT	299
OY	318	GGGGAGAAATATCTGCTTCTCTCAACTTTTCTGTGGCACCATGTCCTCATCTTCTCT	377
Db	300	GGGGAGAAATATCTGCTTCTCTCAACTTTTCTGTGGCACCATGTCCTCATCTTCTCT	359
OY	378	GGAGTGGCTGTGGGCGGTGGTGGATGTTCACCTTGGGGTGTGCCGGCTGGGGGCTCT	437
Db	360	GGAGTGGCTGTGGGCGGTGGTGGATGTTCACCTTGGGGTGTGCCGGCTGGGGGCTCT	419
OY	438	GGAGTTCTTCAGAGCAAACTCAAGTCTCTACCGGAGATATGATCTGCTCAAACTCAT	497
Db	420	GGAGTTCTTCAGAGCAAACTCAAGTCTCTACCGGAGATATGATCTGCTCAAACTCAT	479
OY	498	CGACTCCCTTCAGAAAGCTAACCAAGTCTGTGGCGATATGCGCTG-AAAGCTGGGACC	556
Db	480	CGACTCCCTTCAGAAAGCTAACCAAGTCTGTGGCGATATGCGCTGAAAGCTGGGACC	539
OY	557	TCA-ACGCTACTCTCAATTGACGGGGTGGCAGCTACAGCCGAGAGAAATGCGGGCTCCC	615
Db	540	TCAGACGCTCACTTCAATTGACGGGGTGGCAGCTACAGCCGAGAGAA-TGCGGGGGTCCC	598
OY	616	TTTCTCTGCTCGTGGCCAGATTCGCGGCACAAAAGTTGTGAACAACAAGTGGGATATGAT	675
Db	599	TTTCTCTGCTCGTGGCCAGATTCGCGGCACAAAAGTTGTGAACAACAAGTGGGATATGAT	658
OY	676	GTCAGGATTCAGTGAAGAGCAAGTGGGATGAGTCCATCTTCACGAAAGGTCGATCCAG	735
Db	659	GTCAGGATTCAGTGAAGAGCAAGTGGGATGAGTCCATCTTCACGAAAGGTCGATCCAG	718
OY	736	GCGCTGGAAAGCTGTGGTCCCGCGGAGACATTTACATTGTGGCTGTCTTATCGCCATC	795
Db	719	GCGCTGGAAAGCTGTGGTCCCGCGGAGACATTTACATTGTGGCTGTCTTATCGCCATC	778
OY	796	TCGCTGTTCAGATATTGGCATCTTCCTGCGCAAGGACGCTANTCTCAACATTCAGAGGA	855
Db	779	TCGCTGTTCAGATATTGGCATCTTCCTGCGCAAGGACGCTANTCTCAACATTCAGAGGA	838
OY	856	GTYAAGGCCGGCCATCACTTCTGAGAGCAGATTGAGGAGGCCGAGCTGAGCCACGCTG	915
Db	839	GTYAAGGCCGGCCATCACTTCTGAGAGCAGATTGAGGAGGCCGAGCTGAGCCACGCTG	898
OY	916	GGAGGCCAGAGCCTTTTCTGTCCATACAGCCTTACGTCCAGAGGGAGAGAGCCGACACC	975
Db	899	GGAGGCCAGAGCCTTTTCTGTCCATACAGCCTTACGTCCAGAGGGAGAGAGCCGACACC	958
OY	976	CCAGAGCCAGTGGCCCATCTTAAGCAATACGCGGAGTAGTACGCTCTGTCTTGTGGTCT	1033
Db	959	CCAGAGCCAGTGGCCCATCTTAAGCAATACGCGGAGTAGTACGCTCTGTCTTGTGGTCT	1018
OY	1036	GGTGCTGAAGACCAAGGATCCCGCTTGTATTACGTGCCCAACTTGTGACGTCAATCCCTGT	1095
Db	1019	GGTGCTGAAGACCAAGGATCCCGCTTGTATTACGTGCCCAACTTGTGACGTCAATCCCTGT	1078
OY	1096	GAGTCTTCCACAGACAGAGATATGTCTTTATGTGGAGGTGTGACTCTTGAAGACAGA	1155

Dd	1 079	GAGTCTACCCGAGACGAGAAATGTGTCTTTATGTGGAGTGSTACTCTCGAAGACAGA	1138
Qy	1156	GAGGGCTCTGTGGCTGCCAGGAGGGCTTACTCAGACCCTGCAGCTCAAGCATGTCT	1215
Dd	1139	GAGGGCTCTGTGGCTGCCAGGAGGGCTTGACTACACACCCCCTGCAGCTCAAGCATGTCT	1198
Qy	1216	GCAGGACACCTGTGTCCCCCTCTTCACAATGTGGCWT-CCAGACATCTCTTTGGGTC-ATTC	1272
Dd	1199	GCAGSACACCTGGTCCCTCTCTCCAGTGGCACTCCCAAACATCTCTTTGGGTCATCC	1258
Qy	1273	ACATCTGGGGTNNGCC-----GTGGGTGAGGGAGCCACAGCGTGGACAGGG-----CA	1323
Dd	1259	ACATCTGGGGTGGGCCCCCGGGGTAAAGMAAGAACCCACAGCGCTGGACACAGGGCATCC	1318
Qy	1324	TCTCTCTCATCAAGCAAGCACACATGGGGGGCTTTGCCGTAAACGG	1370
Dd	1319	TCTCTCCATTCACGAAGCAAGCCAGCATGGGGGGCTTCCCGTAACGGG	1365

RESULT 7	
AAA60192	AAA60192 standard; DNA; 2367 BP.
ID	
XX	
AC	AAA60192;
XX	
DT	02-FEB-2001 (first entry)
XX	
DE	Hydrophobic domain protein cDNA HP02956 isolated from KB cells.
XX	
KW	Human; secreted protein; membrane protein; hydrophobic domain;
KW	proliferation control; differentiation induction; material transport;
KW	biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
KW	immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KW	haemostatic; thrombolytic; anti-inflammatory; tumor inhibition;
KW	autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;
KW	gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200029448-A2.
XX	
PD	25-MAY-2000.
XX	
PE	17-NOV-1999; 99MO-JP06412.
XX	
PR	17-NOV-1998; 98JP-0326255.
XX	
PR	22-DEC-1998; 98JP-0364315.
XX	
PR	16-MAR-1998; 99JP-0069811.
XX	
PR	27-APR-1999; 99JP-0119299.
XX	
PR	19-MAY-1999; 99JP-0138169.
XX	
PA	(SAGA ) SAGAMI CHEM RES CENT.
PA	(PROT-) PROTEGENE INC.
XX	
PI	Kato S, Kimura T;
XX	
DR	WPI: 2000-387753/33.
XX	
DR	P-PSDB; AAB12120.
XX	
PT	Proteins comprising hydrophobic regions, such as secretory and membrane
XX	
PT	proteins, useful in research and diagnostics and having various
XX	
PT	activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
XX	
PT	hemostatic, thrombolytic -
XX	
PS	Claim 4; Page 209-211; 410pp; English.
XX	
CC	Secretory proteins play important roles in the proliferation control, the
CC	differentiation induction, the material transport and the biophylaxis of
CC	cells. Membrane proteins have important roles as signal receptors, ion
CC	channels and transporters. The present sequence is the coding sequence
CC	for a human protein which has at least one hydrophobic domain. The
CC	protein encoded by the present sequence may be a secretory or a membrane



CC The invention relates to novel secreted/transmembrane proteins, and  
 CC nucleic acids encoding them. The novel proteins are designated TANGO 339,  
 CC TANGO 353, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO  
 CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a  
 CC murine TANGO 393 is also included within the scope of the invention. The  
 CC invention also encompasses fragments and variants of the proteins of the  
 CC invention, and nucleic acids encoding them. The invention additionally  
 CC relates to host cells comprising a nucleic acid of the invention; methods  
 CC for the production of a protein of the invention; an antibody specific  
 CC for a protein of the invention; methods for detecting a protein or  
 CC nucleic acid of the invention; and methods of identifying agents which  
 CC bind to or modulate the activity of a protein of the invention. The novel  
 CC secreted proteins, nucleic acids encoding them, and antibodies against  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate expression or activity of the secreted  
 CC proteins. The secreted expression of the invention may also be used to  
 CC identify modulators of expression or activity, which may be useful in  
 CC the treatment of disorders associated with the proteins of the  
 CC invention e.g., cancers and immunological disorders. The present  
 CC sequence represents cDNA encoding human TANGO protein of the invention.  
 XN  
 XQ Sequence 813 BP; 163 A; 213 C; 235 G; 202 T; 0 other;

Query Match	58.5%;	Score 811.4;	DB 22;	Length 813;
Best Local Similarity	99.9%;	Pred. No. 1e-203;		
Matches 812; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	67	ATGACTATTATAGTACTCTTAAGCCAAAGTACGCTGCTGGTACAAAGTACTCTTTTC	1220
Dp	1	ATTCCTATTATTATAGTACTCTTAAGCCAAAGTACGCTGCTGGTACAAAGTACTCTTTTC	60
QY	127	ACGTACAACATCATCTTCGTGGTGGATGGATTTGCTCCCTTGGATCGGGCTGGGGCA	186
Dp	61	AGCTACAACATCATCTTCGTGGTGGATGGATTTGCTCCCTTGGATCGGGCTGGGGCA	122
QY	187	TGAGAGCAAAAGGCTGTCTCCGACCTCACCAAGATGACCCGGATGATGAATGCAC	246
Dp	121	TGAGAGCAAAAGGCTGTCTCCGACCTCACCAAGATGACCCGGATGATGAATGCAC	18
QY	247	CGTGTGCTGTGCTCTCATAGTGGGGGGGTATGTTATCACTCGGGGTTCGGCGCTGC	30
Dp	181	CGTGTGCTGTGCTCTCATAGTGGGGGGGTATGTTATCACTCGGGGTTCGGCGCTGC	240
QY	307	GTCGGGGCTCTCGGGACAATATCGCTGCCTCAACTTTTCTGTGGCCATCGTCTC	366
Dp	241	GTCGGGGCTCTCGGGAGAATATCGTCTCTCAACTTTTCTGTGGCCATCGTCTC	30
QY	367	ATTTTCTTCCTGGAGCTGGCTGTGGCCGTGCTGCTCTCTGTTCCAGAGCTGGGTAGG	428
Dp	301	ATTTTCTTCCTGGAGCTGGCTGTGGCCGTGCTGCTCTCTGTTCCAGAGCTGGGTAGG	360
QY	427	GACCGGTTCCGGGAATTTCTTGGAGAGCAACATCAAGCCACCGGAGCATATGATCTG	488
Dp	361	GACCGGTTCCGGGAATTTCTTGGAGAGCAACATCAAGCCACCGGAGCATATGATCTG	420
QY	487	CAAAACCTCATGACTCCCTTCAGAAAGCTAACCACTGCTGTGGGCATATAGGCCCTGAA	548
Dp	421	CAAAACCTCATGACTCCCTTCAGAAAGCTAACCACTGCTGTGGGCATATAGGCCCTGAA	480
QY	547	GACTGGGACCTAACGCTACTTCAATTGCAGCGGTGCCAGCTACAGCCGAGAGAATGC	606
Dp	481	GACTGGGACCTAACGCTACTTCAATTGCAGCGGTGCCAGCTACAGCCGAGAGAATGC	540
QY	607	GGGGTCCCTTCTCTCGTGGGTGGCCAGATCTGCGCAAAAAGTTGTGAACACACAGCT	666
Dp	541	GGGGTCCCTTCTCTCGTGGGTGGCCAGATCTGCGCAAAAAGTTGTGAACACACAGCT	600
QY	667	GGATATGATGATCAGATTACCTCAAGACGAAGTGGGATGAGTCCATTTTCAACGAAAGC	726
Dp	601	GGATATGATGATCAGATTACCTCAAGACGAAGTGGGATGAGTCCATTTTCAACGAAAGC	660
QY	727	TGCATCAGGCGCTGAAAAGTGGCTCCCGCGGAACATTTTACATTTGTGGCTGCTTC	786

Db	661	TGCATCCAGGCGCGCTGGAAAGCTGGCTCCCGCGGAACATTTTACATTTGTGGCTGGCGTCTTC	720
QY	787	ATCGCCATCTCGCTGTTTCAGCATATATTGGCATCTTCTCGCAGAGACGCTGATCTCAGAC	846
Db	721	ATCGCCATCTCGCTGTTTCAGCATATATTGGCATCTTCTCGCAGAGACGCTGATCTCAGAC	780
QY	847	ATCGAGGCAGTGAAGGCCCGGCATCACTTTCGA	879
Db	781	ATCGAGGCAGTGAAGGCCCGGCATCACTTTCGA	813

## RESULT 9

AAE90656  
ID AAE90656 standard; CDNA; 813 BP.

AC AAF90656;

DT 04-MAY-2001 (first entry)

DE Human TANGO 339 K10R variant cDNA ORF, SEQ ID NO:136.

**KW** Secreted protein; transmembrane protein; TANGO; human; drug screening;

KW cytostatic; immunomodulatory; gene therapy; ss.

OS Homo sapiens.

PN WO200109162-A2.

PD 08-FEB-2001.

31-JUL-2000; 2000WO-US20935.

PR 30-JUL-1999; 99US-0365164.

PA (MILL-) MILLENNIUM PHARM INC.

PI Fra

XX

DR P-PSDB; AAB87134.

PT Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful

PT dis

PS Claim 2; Page 279; 332pp; English.

CC The invention relates to novel secreted/transmembrane proteins, and  
CC nucleic acids encoding them. The novel proteins are designated TANGO 339,  
CC TANGO 353, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO  
CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a  
CC murine TANGO 399 is also included within the scope of the invention. The  
CC invention also encompasses fragments and variants of the proteins of the  
CC invention, and nucleic acids encoding them. The invention additionally  
CC relates to host cells comprising a nucleic acid of the invention; methods  
CC for the production of a protein of the invention; an antibody specific  
CC for a protein of the invention; methods for detecting a protein or  
CC nucleic acid of the invention; and methods of identifying agents which  
CC bind to or modulate the activity of a protein of the invention. The novel  
CC secreted proteins, nucleic acids encoding them, and antibodies against  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression or activity of the secreted  
CC proteins. The secreted proteins of the invention may also be used to  
CC identify modulators of expression or activity, which may be useful in  
CC the treatment of disorders associated with the proteins of the  
CC invention e.g., cancers and immunological disorders. The present  
CC sequence represents cDNA encoding human TANGO protein of the invention.

Sequence 813 BP; 162 A; 213 C; 236 G; 202 T; 0 other;

Query Match 58.38; Score 809.8; DB 22; Length 813;





Db	301	ATCTCTCTCTCGAGCTGGCTGTGGCCGCTCTGGCCCTTCTCTTCTTCAGAGACTGGGTAGG	360
QY	427	GACCGGTTCCGGAGTCTTTCGAGAGCAACATCAATCACTTACCGGAGCAGATTCGATCTG	486
Db	361	GACCGGTTCCGGAGTCTTTCGAGAGCAACATCAATCACTTACCGGAGCAGATTCGATCTG	420
QY	487	CAAAACCTCATGAGTCCCTTCAGAAAGCTAACCAAGTGTGGCGCATATGGCCCTGAA	546
Db	421	CAAAACCTCATGAGTCCCTTCAGAAAGCTAACCAAGTGTGGCGCATATGGCCCTGAA	480
QY	547	GACTGGGACCTCAACGCTCACTTCAATTCGACGCGTGCACAGTACACCGCAGAGAGTGC	606
Db	481	GACTGGGACCTCAACGCTCACTTCAATTCGACGCGTGCACAGTACACCGCAGAGAGTGC	540
QY	607	GGGGTCCCTTCTCTGCTGGTGCACAGATCTGCGSCAATAAGTTGTACACACAGTGT	666
Db	541	GGGGTCCCTTCTCTGCTGGTGCACAGATCTGCGSCAATAAGTTGTACACACAGTGT	600
QY	667	GGATATGATGTACAGATTCACGTAAGAAGCAATGGGATGAGTCCATCTTCAGAAAGG	726
Db	601	GGATATGATGTACAGATTCACGTAAGAAGCAATGGGATGAGTCCATCTTCACAAAGG	660
QY	727	TGATCCAGGCGCTGGAAAGCTGGCTCCCGCGAACAATTACATGTGGTGGCGTTC	786
Db	661	TGATCCAGGCGCTGGAAAGCTGGCTCCCGCGAACAATTACATGTGGTGGCGTTC	720
QY	787	ATGCGCATCTCGCTGTTCGCAATATTTTGGCATCTTCTTGGCAAGCAGCTGATCTCAGAC	846
Db	721	ATGCGCATCTCGCTGTTCGCAATATTTTGGCATCTTCTTGGCAAGCAGCTGATCTCAGAC	780
QY	847	ATCGAGCGAGTGAAGCGCGGCATCACTTCTGA	879
Db	781	ATCGAGCGAGTGAAGCGCGGCATCACTTCTGA	813
RESULT 11			
AAAF90658			
ID	AAAF90658	standard; cDNA; 813 BP.	
XX	AAAF90658;		
AC			
XX	04-MAY-2001 (first entry)		
DT			
XX			
DE	Human TANGO 339 A40V variant cDNA ORF, SEQ ID NO:140.		
XX			
KW	Secreted protein; transmembrane protein; TANGO; human; drug screening;		
KW	activity modulator; expression modulator; cancer; immunological disorder		
KW	cytostatic; immunomodulatory; gene therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200109162-A2.		
XX			
PD	08-FEB-2001.		
XX			
PF	31-JUL-2000; 2000WO-US20935.		
XX			
PR	30-JUL-1999; 99US-0365164.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrighton N, Myers PS,		
PI	Pan Y;		
XX			
DR	WPI; 2001-138647/14.		
DR	P-SDB; AAB87136.		
XX			
XX	Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful		
XX	for the prevention, diagnosis and treatment of, e.g. cancers and immune		
XX	disorders.		
XX			
XX	Claim 2; Page 282; 332pp; English		

Query Match	Best Local Similarity	Score	DB	Length
Matches	811; Conservative	0; Mismatches	2; Indels	0; Gaps
67	ATGCACATATTATAGATACCTCTAAAGCCAAAGTCACGCTCTGTAACAGTACCTCTTTTC	58.3%	DB 22:	126
Db	1	ATGCACATATTATAGATACCTCTAAAGCCAAAGTCACGCTCTGTAACAGTACCTCTTTTC		60
QY	127	AGCTACAAACATCATCTTGCTGGTGGAGTGTCTTCCTTTGGAGTCGGGCTGTGGCA		186
Db	61	AGCTACAAACATCATCTTGCTGGTGGAGTGTCTTCCTTTGGAGTCGGGCTGTGGGA		120
QY	187	TGGAGCAAAAGAGTGTCTGCTCCGACTCACCAAAAGTACCCGAGATCATAGATACG		246
Db	121	TGGAGCAAAAGAGTGTCTGCTCCGACTCACCAAAAGTACCCGAGATCATAGATACG		180
QY	247	CCTGTGGGCTGGGCTGCGGATGGTGGGCGTCGATGTTACCTCTGGGCTTCCGGCTGC		306
Db	181	CCTGTGGTGTGCTGCTCGATGGTGGGCGCTGGTGAATTCACCTCTGGGCTTCCGGCTGC		240
QY	307	GTGGGGGCTCTGCGGGAGAAATATCTGCTTGCCTCAACTTTTCTGTGGCACCATCTGCTC		366
Db	241	GTGGGGGCTCTGCGGGAGAAATATCTGCTTGCCTCAACTTTTCTGTGGCACCATCTGCTC		300
QY	367	ATCTTCTCTCGAGAGCTGGGCTGGGCGTGGGCGCTTCTGTTCCAGAGCTGGGTGAG		426
Db	301	ATCTTCTCTCGAGAGCTGGGCTGGGCGCTGGGCGCTTCTGTTCCAGAGCTGGGTGAG		360
QY	427	GACCGGTTCCGGAGATTCTTCAGAGCAACATCAAGTCTTACCGGAGCATATTCATCTG		486
Db	361	GACCGGTTCCGGAGATTCTTCAGAGCAACATCAAGTCTTACCGGAGCATATTCATCTG		420
QY	487	CAAAACCTCATCACTCCCTTGAAGAGTATACAGTCTGTGGCGCATATGGCCCTGAA		546
Db	421	CAAAACCTCATCACTCCCTTGAAGAGTATACAGTCTGTGGCGCATATGGCCCTGAA		480
QY	547	GACTGGGACCTCAAGCTCTACTTCAATTGACAGCGGTGTGCCACTTACGCCAGAGAAAGTC		606
Db	481	GACTGGGACCTCAAGCTCTACTTCAATTGACAGCGGTGTGCCACTTACGCCAGAGAAAGTC		540
QY	607	GGGGTCCCTTCCTTCCTGCTGGCGACATCCTCTGGCGAAAAGTTTGAGCACACAGTGT		666
Db	541	GGGGTCCCTTCCTTCCTGCTGGCGACATCCTCTGGCGAAAAGTTTGAGCACACAGTGT		600
QY	667	GGATATGATGTCAAGATTACGTAAGAGCAAGTGGATGAGTCCATCTTCACGAAAGGC		726
Db	601	GGATATGATGTCAAGATTACGTAAGAGCAAGTGGATGAGTCCATCTTCACGAAAGGC		660
QY	727	TGCATCCAGGCGCTGGAAAGCTGGCTCCCGCGGAAACATTTTCATTGTGGCTGGCGATTC		786



|||||  
Db 661 TGCATCCAGCGCCGTGAAAGCTGCCCGCGAGACATTTACATTTGTGGCTGGCGCTCTTC 720  
QY 787 ATCGGCATCTGCGTGTGGAGATTTTGGCATCTCTCCGCGAGAGACCGCTATCTCAAC 846  
Db 721 ATCGGCATCTGCGTGTGGAGATTTTGGCATCTCTCCGCGAGAGACCGCTATCTCAAC 780  
QY 847 ATCGAGGAGTGAAGCGCGGCATCATCTTCTGA 879  
Db 781 ATCGAGGAGTGAAGCGCGGCATCATCTTCTGA 813  
  
RESULT 12  
AAF90659  
ID AAF90659 standard; cDNA; 813 BP.  
XX  
AC AAF90659;  
XX  
DT 04-MAY-2001 (first entry)  
XX  
XX Human TANGO 339 D60E variant cDNA ORF, SEQ ID NO:142.  
DE  
XX  
XX Secreted protein; transmembrane protein; TANGO; human; drug screening;  
KW activity modulator; expression modulator; cancer; immunological disorder;  
KW cytosolic; immunomodulatory; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200109162-A2.  
XX  
XX 08-FEB-2001.  
XX  
XX 31-JUL-2000; 2000MO-US20935.  
XX  
XX 30-JUL-1999; 9905-0365164.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrighton N, Myers PS;  
PI Pan Y;  
XX  
XX WPI: 2001-138647/14.  
XX  
XX P-PSDB; AAB87137.  
XX  
XX Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful  
PT for the prevention, diagnosis and treatment of, e.g. cancers and immune  
PT disorders -  
XX  
XX  
XX Claim 2; Page 283; 332pp; English.  
XX  
XX The invention relates to novel secreted/transmembrane proteins, and  
CC nucleic acids encoding them. The novel proteins are designated TANGO 339,  
CC TANGO 353, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO  
CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a  
CC murine TANGO 393 is also included within the scope of the invention. The  
CC invention also encompasses fragments and variants of the proteins of the  
CC invention, and nucleic acids encoding them. The invention additionally  
CC relates to host cells comprising a nucleic acid of the invention; methods  
CC for the production of a protein of the invention; an antibody specific  
CC for a protein of the invention; methods for detecting a protein or  
CC nucleic acid of the invention; and methods for identifying agents which  
CC bind to or modulate the activity of a protein of the invention. The novel  
CC secreted proteins, nucleic acids encoding them, and antibodies against  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression or activity of the secreted  
CC proteins. The secreted proteins of the invention may also be used to  
CC identify modulators of expression or activity, which may be useful in  
CC the treatment of disorders associated with the proteins of the  
CC invention e.g., cancers and immunological disorders. The present  
CC sequence represents cDNA encoding human TANGO protein of the invention.  
XX  
XX Sequence 813 BP; 163 A; 212 C; 236 G; 202 T; 0 other;

Query Match 58.3%; Score 809.8; DB 22; Length 813;  
Best Local Similarity 99.8%; Pred. No. 2,7e-203;  
Matches 811; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 67 ATGCATATTATATAGTACTCTAACGCCAAGTACGTCTGTGTACAGTACTCTTTC 126  
Db 1 ATGCATATTATATAGTACTCTAACGCCAAGTACGTCTGTGTACAGTACTCTTTC 60  
QY 127 AGTACAAATCATCTCTGTGGTGGAGTGTCTCTTGTGGAGTGGGCTGTGGGCA 186  
Db 61 AGTACAAATCATCTCTGTGGTGGAGTGTCTCTTGTGGAGTGGGCTGTGGGCA 120  
QY 187 TGGAGCCAAAAGGCTGTCTGTCCGACCTCACCAAGTACCCGAGTATGGAATGAC 246  
Db 121 TGGAGCCAAAAGGCTGTCTGTCCGACCTCACCAAGTACCCGAGTATGGAATGAC 180  
QY 247 CCTGTGTGCTGTGTCTGTATGTGTGGGCTGTGTATTCACCCGTGGGCTGTGC 306  
Db 181 CCTGTGTGCTGTGTCTGTATGTGTGGGCTGTGTATTCACCCGTGGGCTGTGC 240  
QY 307 GTGGGGGCTGTGGGGGAGATATCTGTCTCACTTTCTGTGGACATCGTGTGC 366  
Db 241 GTGGGGGCTGTGGGGGAGATATCTGTCTCACTTTCTGTGGACATCGTGTGC 300  
QY 367 ATCTTCTTCCGTGAGCTGTGGCTGTGGCGCTTCCCTGTCCAGACCTGGGTAGG 426  
Db 301 ATCTTCTTCCGTGAGCTGTGGCTGTGGCGCTTCCCTGTCCAGACCTGGGTAGG 360  
QY 427 GACCGGTTCCGGGAGTCTTTCGAGAGCAACATCAATCTTACCGGAGCATATGATCTG 486  
Db 361 GACCGGTTCCGGGAGTCTTTCGAGAGCAACATCAATCTTACCGGAGCATATGATCTG 420  
QY 487 CAAAACCTCATGCACTCCCTTCGAAAGCTAACAGTGTGGCGATATGGCCCTGAA 546  
Db 421 CAAAACCTCATGCACTCCCTTCGAAAGCTAACAGTGTGGCGATATGGCCCTGAA 480  
QY 547 GACTGGACCTCAACGCTACTTCAATTGGACGGGTGCCAGCTACAGCCGAGAAATGC 606  
Db 481 GACTGGACCTCAACGCTACTTCAATTGGACGGGTGCCAGCTACAGCCGAGAAATGC 540  
QY 607 GGGGTCCCTTCTCCTGTCTGTGCGAGATCTGCGCAAAAAGTTGTGAACACAGTGT 666  
Db 541 GGGGTCCCTTCTCCTGTCTGTGCGAGATCTGCGCAAAAAGTTGTGAACACAGTGT 600  
QY 667 GGATATGATGTCAAGATTTACAGTGAAGCAAGTGGGATATGATTCATCTTCAGAAAGGC 726  
Db 601 GGATATGATGTCAAGATTTACAGTGAAGCAAGTGGGATATGATTCATCTTCAGAAAGGC 660  
QY 727 TGCATCCAGGCGCTGGAAGCTGTGCCCGGGAACATTTACATTTGGCTGGCGTCTTC 786  
Db 661 TGCATCCAGGCGCTGGAAGCTGTGCCCGGGAACATTTACATTTGGCTGGCGTCTTC 720  
QY 787 ATCGGCATCTGCGTGTGGAGATTTTGGCATCTTCTCCGCGAGAGCGTGTATCTCAGAC 846  
Db 721 ATCGGCATCTGCGTGTGGAGATTTTGGCATCTTCTCCGCGAGAGCGTGTATCTCAGAC 780  
QY 847 ATCGAGGAGTGAAGCGCGGCATCATCTTCTGA 879  
Db 781 ATCGAGGAGTGAAGCGCGGCATCATCTTCTGA 813  
  
RESULT 13  
AAA93675  
ID AAA93675 standard; DNA; 816 BP.  
XX  
XX AAA93675;  
XX  
XX 16-JAN-2001 (first entry)  
XX  
XX Human tetraspanin-like protein splice variant SECX 3911675-2 DNA.  
XX  
XX SECX protein; human; secreted; membrane-associated; cancer;  
KW proliferation regulator; differentiation regulator; non-malignant tumour;





CC and thrombolytic activity, anti-inflammatory activity and tumour  
inhibition activity. The present sequence could therefore be used for  
treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
disease, and cancer via gene therapy.

XX  
SQ Sequence 639 BP; 131 A; 163 C; 188 G; 157 T; 0 other;

Query Match 32.9%; Score 456.4; DB 21; Length 639;  
Best Local Similarity 78.8%; Pred. No. 3.3e-110;  
Matches 638; Conservative 0; Mismatches 1; Indels 171; Gaps 1;

```
QY 67 ATGCACATATATAGTACTTAAGCCCAAGTCACTGCTGTGCAAGTACCTCTTTC 126
    |||||||
Db 1 ATGCACATATATAGTACTTAAGCCCAAGTCACTGCTGTGCAAGTACCTCTTTC 60

QY 127 ACCTACACATCATCTTCTGCTGCTGAGTGTCTTCTTCTGAGTGGGCTGTGGCA 186
    |||||||
Db 61 ACCTACACATCATCTTCTGCTGCTGAGTGTCTTCTTCTGAGTGGGCTGTGGCA 120

QY 187 TGGAGCGAAAGGGTGTCTGTCCGACCTACCAAGTGAACCGGATGATGGAATCGAC 246
    |||||||
Db 121 TGGAGCGAAAGGGTGTCTGTCCGACCTACCAAGTGAACCGGATGATGGAATCGAC 180

QY 247 CCGTGGTGTCTGCTGATGGTGGGCTGTGATGTTCAACCGGGGTTCCGGGCTGC 306
    |||||||
Db 181 CCGTGGTGTCTGCTGATGGTGGGCTGTGATGTTCAACCGGGGTTCCGGGCTGC 240

QY 307 GTGGGGGCTCTGCGGAGAAATATCTGCTGCTCAACTTTTCTGTGGACCAATCGTCTC 366
    |||||||
Db 241 GTGGGGGCTCTGCGGAGAAATATCTGCTGCTCAACTT----- 278

QY 367 ATCTTCTTCTGAGCTGGCTGTGGCGGTGCTGCTTCCGTTCCAGACTGGTGAGG 426
    |||||||
Db 279 ----- 278

QY 427 GACCGGTTCCGGAGTCTTCGAGACAATCAAGTCTACCGGGAGATATCGATCTG 486
    |||||||
Db 279 ----- 278

QY 487 CAAAACCTATGACTCCCTTCAGAAAGCTAACCAAGTGTGTGGGCAATATGGCCCTGAA 546
    |||||||
Db 279 -----TAACCAAGTGTGTGTGGGCAATATGGCCCTGAA 309

QY 547 GACTGGACCTCAAGTCTACTTCAATTGCAAGCGGTGCCAGCTACAGCCGAGAGAAAGTC 606
    |||||||
Db 310 GACTGGACCTCAAGTCTACTTCAATTGCAAGCGGTGCCAGCTACAGCCGAGAGAAAGTC 369

QY 607 GGGGTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
    |||||||
Db 370 GGGGTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429

QY 667 GGATATGATGTAGATTCAGTGAAGAGCAAGTGGGATGAGTCCATCTTCAGAAAGGC 726
    |||||||
Db 430 GGATATGATGTAGATTCAGTGAAGAGCAAGTGGGATGAGTCCATCTTCAGAAAGGC 489

QY 727 TGCATCCAGAGCGCTGGAAGCTGCTCCCGGGAACATTTACATTTGCTGCTGCTGCTG 786
    |||||||
Db 490 TGCATCCAGAGCGCTGGAAGCTGCTCCCGGGAACATTTACATTTGCTGCTGCTGCTG 549

QY 787 ATGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
    |||||||
Db 550 ATGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609

QY 847 ATCGAGGCAAGTGAAGCGCGGCATCATCTTC 876
    |||||||
Db 610 ATCGAGGCAAGTGAAGCGCGGCATCATCTTC 639
```

Search completed: December 19, 2002, 03:59:34  
Job time : 268 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## SUMMARIES

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 19, 2002, 05:19:27 : Search time 259 Seconds  
(without alignments)  
2347.643 Million cell updates/sec

Title: US-09-905-674-2  
Perfect score: 1449  
Sequence: 1 MHYRYSNAKYSWCWKYLLF.....IFLARTLISDIEAVKAGHHF 270

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO/US0995674/runat\_16122002.111845\_6745/app.query.fasta\_1.455  
-DB=N.Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptco -NOR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0995674 @cgn.1.1.125 @runat.16122002.111845\_6745 -NCPU=6 -ICPU=3  
-NO\_XIPY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: N.Geneseq\_101002:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1449	100.0	813	22	AAF90629	Human TANGO 339 CD
2	1449	100.0	1388	24	AA29623	Human tetraspan pr
3	1449	100.0	2715	22	AAF90628	Human TANGO 339 CD
4	1446	99.8	813	22	AAF90656	Human TANGO 339 K1
5	1446	99.8	813	22	AAF90657	Human TANGO 339 F2
6	1445	99.7	813	22	AAF90658	Human TANGO 339 A4
7	1445	99.7	813	22	AAF90659	Human TANGO 339 D6
8	1436	98.1	2672	21	AA65259	Human secreted pro
9	1425	98.3	2623	23	AA592016	DNA encoding novel
10	1419	97.9	1428	21	AA93623	Human tetraspanin-
11	1375	94.9	1667	21	AA65350	Human secreted pro
12	1120.5	77.3	639	21	AA60182	Hydrophobic domain
13	1120.5	77.3	2367	21	AAA60192	Human tetraspanin-
14	1049	72.4	816	21	AA93675	Human tetraspanin-
15	936	64.6	601	23	AA80667	DNA encoding novel
16	867	59.8	1655	21	AA98195	Human colon cancer
17	858	59.2	1126	22	AA90014	Clone HMTMH27 codi
18	858	59.2	2091	21	AA96482	CDNA encoding a hu
19	858	59.2	3185	24	AA172633	Human 23228 CDNA.
20	850	58.7	1644	24	AB190838	Human polynucleoti
21	841.5	58.1	2538	24	AA21883	Human TM4SF recept
22	829	57.2	1174	22	AA90015	Clone HE8EJ16 codi
23	820.5	56.6	1178	22	AA90020	Clone HE8EJ16 codi
24	729.5	50.3	1110	22	AA20854	Polynucleotide seq
25	729.5	50.3	1110	22	AA59291	Human CDNA encodin
26	729.5	50.3	1110	24	ABA90960	Human polynucleoti
27	720	49.7	1988	24	ABK35735	CDNA sequence #126
28	681.5	47.0	1214	22	AA58068	DNA encoding novel
29	568	39.2	852	22	AA54553	Human CDNA encodin
30	568	39.2	1932	21	AA298047	Human secreted pro
31	568	39.2	1932	22	AA21660	Human secreted pro
32	568	39.2	1932	24	ABK69756	Human secreted pro
33	567	39.1	864	24	AA172287	NEF-4 antisense mo
34	557	38.4	810	23	AB114835	Drosophila melanog
35	550	38.0	2638	24	AB190058	Human polynucleoti
36	541	37.3	1187	23	AB107929	Drosophila melanog
37	522.5	36.1	3244	23	AB107928	Drosophila melanog
38	510	35.2	1120	23	AB116799	Drosophila melanog
39	508	35.1	3009	23	AB114834	Drosophila melanog
40	485	33.5	839	23	AAH34927	Human colon cancer
41	433	29.9	3303	23	AB116798	Drosophila melanog
42	412.5	28.5	1725	22	AA90017	Clone HOFAD65 codi
43	404.5	27.9	1694	19	AAV33190	Secreted protein B
44	403.5	27.8	1772	20	AA200457	Human secreted pro
45	396	27.3	1749	22	AA90023	Clone HOFAD65 codi

## ALIGNMENTS

RESULT 1  
AAF90629  
ID AAF90629 standard; CDNA; 813 BP.

AC AAF90629;  
XX  
XX 04-MAY-2001 (first entry)

DE Human TANGO 339 cDNA ORF, SEQ ID NO:2.

KX Secreted protein; transmembrane protein; TANGO; human; drug screening;  
KW activity modulator; expression modulator; cancer; immunological disorder;  
KW cytosolic; immunomodulatory; gene therapy; ss.

OS Homo sapiens.  
XX  
XX  
XX WC200109162-A2.



DR P-PSDB: AAE18535.  
 XX New inhibitor of tetraspan protein useful for reducing the expression  
 PT of activities of tetraspan protein in cells, and for treating a  
 PT hyperproliferative disorder, especially cancer in a mammal  
 XX  
 XX Claim 2; Page 84-85; 88pp; English.  
 XX  
 CC The invention relates to inhibitors of tetraspan protein, TSPAN-7  
 CC which are used to reduce or decrease the expression of TSPAN-7  
 CC in a mammalian cell and for treating hyperproliferative disorder,  
 CC especially cancer in a mammal. The invention also provides TSPAN-7  
 CC polypeptides and polynucleotides. TSPAN-7 inhibitors are useful for  
 CC inhibiting proliferation of mammalian cells, including tumor cells,  
 CC for decreasing the side effects of cancer therapy and for treating  
 CC neoplastic diseases. They are used to modulate TSPAN-7 expression  
 CC and function in cancer cells, particularly in prostate cancer cells.  
 CC They are also used in antisense therapy. The present sequence is  
 CC human TSPAN-7 protein encoding cDNA.  
 CC  
 XX  
 SQ Sequence 1388 BP; 270 A; 390 C; 411 G; 313 T; 4 other;  
 Alignment Scores:  
 Pred. No.: 1,24e-155 Length: 1388  
 Score: 1449.00 Matches: 270  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-905-674-2 (1-270) x AAD29623 (1-1388)  
 QY 1 METHISTYRARGTYSERASNAALYSVALISERCYSTRTPTRYLYSTYRLEUENPHE 20  
 DB 67 ATGCACATATATAGATACCTCTAACCCCAAGTCACCTGCTGTACCAAGTACCTCTTTC 126  
 QY 21 SERTYRASNILEIAPHETRPLEUALAGLYVALPHELEGLYVALGLYLEUTRPA 40  
 DB 127 ACCTACACATCATCTCTGCTGGTGGCTGAGTTGCTCTTCCCTGGAGTGGGGCTGGGCA 186  
 QY 41 TRPSERGLULYSGLYVALLEUSERASPLEUTHRLYSVALTHRMETHISGLYLEASP 60  
 DB 187 TGGAGCGAAAGGGTGTGTCTCCACCTCACCAAGTACCCGGATCATGAAATCAGC 246  
 QY 61 PROVALLEULEUVALLEUMETVALGLYVALMETPHEHLEUGLYPHALAGLYYS 80  
 DB 247 CCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306  
 QY 81 VALGLYALALEUALRGUASNILECYSLEULUANPHEPHECYGLYTHRIIEVALLEU 100  
 DB 307 GTGGGGGCTCTCGGGAGAAATATGCTGCTCAACTTTTCTGTGGCACCATCTGCTC 366  
 QY 101 IIEPHEPHELEUGLULEUALVALALAVALLLEUALIAPHELEUPHGLINASPTRYVALRG 120  
 DB 367 AACTCTCTTCTGAGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAGG 426  
 QY 121 ASPATGPHEARGLUPHEPHEGLUSERASNILEYSSERTYRARGASAPNILEASPLEU 140  
 DB 427 GACCGGCTCCGGAGATTCTTCAGAGCAACATCAAGTCTTCCGGGAGATATCATCTG 486  
 QY 141 GLINASLEULEASPSERLEUGLINSALASNGINCYSYSGLYALATYRGLYPROGLU 160  
 DB 487 CAAACCTCATGACTCCCTTCAGAAAGCTAACCAAGTCTGTGGGCGCATATGGCCCTAA 546  
 QY 161 ASPTTPASPLEUASNAVALLYRPHASNGYSSERGLYALASERTYRSEARGLULYSYS 180  
 DB 547 GACTGGAGCTTCAACGCTACTTCAATTCACGCGGTGCACATACAGCGAGAGAGTGC 606  
 QY 181 GLYVALPROPHESERCYSYVALPROASPPROALAGLINSVALVALASPTHGLININS 200  
 DB 607 GGGGGTCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666  
 QY 201 GLYTYRASPVALARGILEGLINLEULYSERYSTRPASPGIUSERLIEPHEHRLYSGLY 220

DB 667 GGATATGATGTCACGATTACAGTCAAGAGCAAGGGGATGATGCTCATCTTCAGAAAGGC 726  
 QY 221 CYSILLEGINALALEUGLUSERTRIPLEUPROARGASNILEYRILEVALALAGLYVALPHE 240  
 DB 727 TGCATCCAGGCGCTGGAAAGCTGCTCCCGGAAACATTACATTTGCTGCTGCTTTC 786  
 QY 241 ILEALATIESERLEULEUGLINSLEPHEGLYIIEPHEULALARGTHRLIEULEISERASP 260  
 DB 787 ATGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846  
 QY 261 ILEGLUALVALYLSALAGLYHSHISHPHE 270  
 DB 847 ATCAGAGGAGTGAAGGCGCGGCATCACTTC 876  
 RESULT 3  
 ID AAF90628  
 XX AAF90628 standard; cDNA; 2715 BP.  
 AC AAF90628;  
 XX  
 DT 04-MAY-2001 (first entry)  
 XX  
 DE Human TANGO 339 cDNA, SEQ ID NO:1.  
 XX  
 KW Secreted protein; transmembrane protein; TANGO; human; drug screening;  
 KW activity modulator; expression modulator; cancer; immunological disorder;  
 KW cyostatic; immunomodulatory; gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200109162-A2  
 XX  
 PD 08-FEB-2001.  
 XX  
 PE 31-JUL-2000; 2000WO-US20935.  
 XX  
 PF 30-JUL-1999; 99US-0365164.  
 PR  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrighton N, Myers PS;  
 PI Pan Y;  
 PI  
 XX WPI: 2001-138647/14.  
 DR P-PSDB: AAB87034, AAB87035, AAB87036.  
 DR  
 XX  
 PS  
 PS Claim 2; Page 219-220; 332pp; English.  
 XX  
 CC The invention relates to novel secreted/transmembrane proteins, and  
 CC nucleic acids encoding them. The novel proteins are designated TANGO 339,  
 CC TANGO 353, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO  
 CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a  
 CC murine TANGO 393 is also included within the scope of the invention. The  
 CC invention also encompasses fragments and variants of the proteins of the  
 CC invention, and nucleic acids encoding them. The invention additionally  
 CC relates to host cells comprising a nucleic acid of the invention; methods  
 CC for the production of a protein of the invention; an antibody specific  
 CC for a protein of the invention; methods for detecting a protein or  
 CC nucleic acid of the invention; and methods of identifying agents which  
 CC bind to or modulate the activity of a protein of the invention. The novel  
 CC secreted proteins, nucleic acids encoding them, and antibodies against  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate expression or activity of the secreted  
 CC proteins. The secreted proteins of the invention may also be used to  
 CC identify modulators of expression or activity, which may be useful in  
 CC the treatment of disorders associated with the proteins of the  
 CC invention e.g., cancers and immunological disorders. The present  
 CC sequence represents cDNA encoding human TANGO protein of the invention.

```

XX      Sequence 2715 BP: 559 A; 760 C; 715 G; 681 T; 0 other:
SQ
Alignment Scores:
Pred. No.:      3,25e-155      Length:      2715
Score:          1449.00      Matches:      270
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             22      Gaps:      0

US-09-905-674-2 (1-270) x AAF90628 (1-2715)

QY      1 MethISTYrTYrArGTySerASnaIAlaYSaISerCYSTrPTyRlySTyRleuLeuphe 20
      |||||||
DB      210 ATGCACATATTATNGATACCTTAACGCCAAGCTACGCTGCTGATACAGTACTCTTTTC 269

QY      21 SerTYrASnIleIlePheTrPleuAaGlyValAlPheLeuGlyValGlyLeuTrpAla 40
      |||||||
DB      270 AGCTACACATCATCTTCTGTGGCTGGAGTTGCTCTTCCCTTGAGTGGGCTGTGGGCA 329

QY      41 TrpSerGluLysGlyValLeuSerAspLeuTrpLysValThrArgMethISgIlyLeasp 60
      |||||||
DB      330 TGGAGCGAAAGGGTGTGTCTCCGACCTCCACCAAGTACCCGGATGATGAAATCGAC 389

QY      61 ProValIleuValLeuMetValGlyValAlMetPheThrLeuGlyPheAlaGlyCys 80
      |||||||
DB      390 CCGTGTGCTGTGCTCTGATGAGGCGCTGTGATGTTCCACCTGGGGTGGCGGCTGTC 449

QY      81 ValGlyAlaLeuArgGluASnIleCYsLeuLeuASnPhenPheCYsGlyThrIleValleu 100
      |||||||
DB      450 GTGGGGGCTCTGGGGAGATATCTGCTCAACTTTTCTGTGGCCACATCGTGTCTC 509

QY      101 IlePhePheLeuGluLeuAlaValAlaValIleuAlaPheLeuPheGluAspTrpAlaArg 120
      |||||||
DB      510 ATCTTTCTCTGGAGCTGTGCTGTGGCCGTGTGGCTTCTCTTCCAGAGACTGGGTGAGG 569

QY      121 AspArgPheArgGluPhePheGluSerASnIleLYsSerTYrArGAspApIleAspLeu 140
      |||||||
DB      570 GACCGTTCGGGAGTCTTCCAGAGCAACATCAAGTCTACCGGGAGCATTCGATCTG 629

QY      141 GluASnLeuIleAspSerLeuGlnLYsAlaASnGlnCYsCYsGlyAlaTrpGlyProGlu 160
      |||||||
DB      630 CAAAACTCATGCACTCCCTCAGAAAGCTAACCAAGTCTGTGGCGCATATGCGCCCTGAA 689

QY      161 AspTrpAspLeuASnValIyrPheASnCYsSerGlyAlaSerTYrSerArgGluLysCys 180
      |||||||
DB      690 GATGGGACCTCAACGCTTACTTCAATTCACGGGTGCGACGCTACAGCCGAGAGAAGTGC 749

QY      181 GlyValProPheSerCYsCYsValProAspProAlaGlnLYsValAlaASnThrGlnCYs 200
      |||||||
DB      750 GGGGTCCTCTCCCTCGTGGCGGCAAGATCCTGGCGCAAAAGTTGTGAACACACAGTGT 809

QY      201 GlyTYrAspValArgIleGlnLeuLYsSerLYSTrPASPgluSerIlePheThrIlysgly 220
      |||||||
DB      810 GGATATGATGTCTGGATTTCAGCTGAAGACCAAGTGGATGATGCATCTTCGGAAGGCG 869

QY      221 CysIleGlnAlaLeuGluSerTrPleuProArgASnIleTYrIleValAlaGlyValPhe 240
      |||||||
DB      870 TGCATTCAGGCGCTGGAAGCTGGCTCCCGCGGACATTTACATTTGTGGCTGTCTTC 929

QY      241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerASP 260
      |||||||
DB      930 ATCGCATCTCGCTGTGTGAGATATTGGCATCTTCTCGCAAGACGCGTGTATCTCAGAC 989

QY      261 IleGluAlaValLYsAlaGlyIleHisIshPhe 270
      |||||||
DB      990 ATCGAGGCAGTGAAGCCGCGCATCTTC 1019

RESULT 4
AAF90656
ID      AAF90656 standard; cDNA; 813 BP.
XX

```

```

AC      AAF90656;
XX
XX      04-MAY-2001 (first entry)
DE      Human TANGO 339 K10R variant cDNA ORF, SEQ ID NO:136.
XX
XX      Secreted protein; transmembrane protein; TANGO; human; drug screening;
KW      activity modulator; expression modulator; cancer; immunological disorder;
KW      cytostatic; immunomodulatory; gene therapy; ss.
XX
XX      Homo sapiens.
OS
XX
XX      W0200109162-A2.
PN
XX
XX      08-FEB-2001.
PD
XX
XX      31-JUL-2000; 2000WO-US20935.
FE
XX
XX      30-JUL-1999; 99US-0365164.
PR
XX
XX      (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX      Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrigton N, Myers PS;
PI      Pan Y;
XX
XX      WPI: 2001-138647/14.
DR
XX
XX      P-PSDB: AAB87134.
PT
XX
XX      Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
PT      for the prevention, diagnosis and treatment of, e.g. cancers and immune
PT      disorders -
XX
PS      Claim 2; Page 279; 332pp; English.
XX
XX      The invention relates to novel secreted/transmembrane proteins, and
CC      nucleic acids encoding them. The novel proteins are designated TANGO 339,
CC      TANGO 353, TANGO 358, TANGO 365, TANGO 369, TANGO 383, TANGO
CC      393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
CC      murine TANGO 393 is also included within the scope of the invention. The
CC      invention also encompasses fragments and variants of the proteins of the
CC      invention, and nucleic acids encoding them. The invention additionally
CC      relates to host cells comprising a nucleic acid of the invention; methods
CC      for the production of a protein of the invention; an antibody specific
CC      for a protein of the invention; methods for detecting a protein or
CC      nucleic acid of the invention; and methods of identifying agents which
CC      bind to or modulate the activity of a protein of the invention. The novel
CC      secreted proteins, nucleic acids encoding them, and antibodies against
CC      them may be used in the prevention, diagnosis and treatment of diseases
CC      associated with inappropriate expression or activity of the secreted
CC      proteins. The secreted proteins of the invention may also be used to
CC      identify modulators of expression or activity, which may be useful in
CC      the treatment of disorders associated with the proteins of the
CC      invention e.g., cancers and immunological disorders. The present
CC      sequence represents cDNA encoding human TANGO protein of the invention.
XX
XX      Sequence 813 BP; 162 A; 213 C; 236 G; 202 T; 0 other:
SQ
Alignment Scores:
Pred. No.:      1,26e-155      Length:      813
Score:          1446.00      Matches:      269
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.63%      Mismatches: 0
Query Match:    99.79%      Indels:      0
DB:             22      Gaps:      0

US-09-905-674-2 (1-270) x AAF90656 (1-813)

QY      1 MethISTYrTYrArGTySerASnaIAlaYSaISerCYSTrPTyRlySTyRleuLeuphe 20
      |||||||
DB      1 ATGCACATATTATNGATACCTTAACGCCAAGCTACGCTGCTGATACAGTACTCTTTTC 60

QY      21 SerTYrASnIleIlePheTrPleuAaGlyValAlPheLeuGlyValGlyLeuTrpAla 40
      |||||||

```



Db 61 AGCTACACATCATCTTGGTGGCTGGAGTTGCTTCCTTGAGCTCGGAGCTGGGCA 120  
 QY 41 TTPSerGIuLysGIyValLeuSerAspLeuThrLyValThrArgMetHisGIyIleAsp 60  
 Db 121 TGGAGGAAAGGGTGTGCTGTCCGACCTCACCAGAGTACCCGGATGGATGATGATCGAC 180  
 QY 61 ProValValLeuValLeuMetValGIyValValMetPheThrLeuGIyPheAgiLys 80  
 Db 181 CCTGTGTGCTGTGCTGTGATGGGCGGTGATGTTCACCTGGGGTTCGCCGGCTGC 240  
 QY 81 ValGIyAlaLeuArgLysnIleCysLeuLeuAsnPhenPheCysGIyThrIleValLeu 100  
 Db 241 GTGGGGGCTCTGGGAGATATCTGCTGCTCAACTTTTGTGTGGCCATCTGCTGC 300  
 QY 101 IlePhePheLeuGIuLeuAlaValAlaValLeuAlaPheLeuPheGlnAspIrrValArg 120  
 Db 301 ATCTTCTCTGAGCTGTGCTGTGGCCGTGCTGGCTTCTTCTTCAGAGCATGGGTGAGG 360  
 QY 121 AspArgPheArgLysPhePheGlnSerAsnIleLysSerTyrArgAspAspIleAspLeu 140  
 Db 361 GACCCGCTTCGGGAGATTCTTCAGAGACACATCAAGTCTACCGGGAGCATTCGATCG 420  
 QY 141 GlnAsnLeuIleAspSerLeuGIuLysAlaAsnGIuLysCysGIyAlaIrrGIyProGIu 160  
 Db 421 CAAACCTCATGACTCCCTTCAAGAAAGCTAACCAAGTGTGGCGCATATGGCCCTGAA 480  
 QY 161 AspIrrAspLeuAsnValTyrPheAsnCysSerGIyAlaSerTyrSerArgLysCys 180  
 Db 481 GACTGGAGCTCAAGCTCTACTTCAATTGCACGGGTGCAGTACAGCCGAGAGAAAGTGC 540  
 QY 181 GIyValIrrPheSerCysCysValProAspProAlaGIuLysValValAsnThrGIuLys 200  
 Db 541 GGGGTCCCTCTCTCTGCTGCTGCAGATCTCTGCGCAAAAGTGTGTAACACACAGTGT 600  
 QY 201 GIyTyrAspValArgIleGlnLeuLysSerLysTrrPaspLysSerIlePheThrIrrGIy 220  
 Db 601 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
 QY 221 CysIleGlnAlaLeuGIuSerTrrPleuProArgAsnIleTyrIleValAlaGIyValPhe 240  
 Db 661 TGCATCCAGGCGCTGGAAGCTGGCTCCCGCGGACATTTCATATGTGCTGGCTCTTC 720  
 QY 241 IleAlaIleSerLeuLeuGIuIlePheGIyIlePheLeuAlaArgThrLeuIleSerAsp 260  
 Db 721 ATCGGCATCTCGCTGTGAGATATTTGGCATCTCTCGCAAGACGCTGATCTCAGAC 780  
 QY 261 IleGluAlaValIleAlaGIyHisHisPhe 270  
 Db 781 ATCGAGGCAGTGAAGCGCGCCATCATCTTC 810  
 RESULT 5  
 AAF90657  
 ID AAF90657 standard; cDNA: 813 BP.  
 AC AAF90657;  
 XX  
 DT 04-MAY-2001 (first entry)  
 XX  
 DE Human TANGO 339 F20Y variant cDNA ORF, SEQ ID NO:138.  
 XX  
 KW Secreted protein; transmembrane protein; TANGO; human; drug screening;  
 KW activity modulator; expression modulator; cancer; immunological disorder;  
 KW cytosolic; immunomodulatory; gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200109162-A2.  
 XX  
 PD 08-FEB-2001.  
 XX  
 PF 31-JUL-2000; 2000WO-US20935.  
 XX  
 PR 30-JUL-1999; 99US-0365164.

XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrighton N, Myers PS;  
 PI Pan Y;  
 XX  
 DR WPI; 2001-138647/14.  
 XX  
 PT P-PSDB; AAB87135.  
 PT  
 PS Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful  
 PS for the prevention, diagnosis and treatment of, e.g. cancers and immune  
 PS disorders -  
 PS  
 PS Claim 2; Page 280-281; 332pp; English.  
 CC  
 CC The invention relates to novel secreted/transmembrane proteins, and  
 CC nucleic acids encoding them. The novel proteins are designated TANGO 339,  
 CC TANGO 353, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO  
 CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a  
 CC murine TANGO 393 is also included within the scope of the invention. The  
 CC invention also encompasses fragments and variants of the proteins of the  
 CC invention, and nucleic acids encoding them. The invention additionally  
 CC relates to host cells comprising a nucleic acid of the invention; methods  
 CC for the production of a protein of the invention; an antibody specific  
 CC for a protein of the invention; methods for detecting a protein or  
 CC nucleic acid of the invention; and methods of identifying agents which  
 CC bind to or modulate the activity of a protein of the invention. The novel  
 CC secreted proteins, nucleic acids encoding them, and antibodies against  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate expression or activity of the secreted  
 CC proteins. The secreted proteins of the invention may also be used to  
 CC identify modulators of expression or activity, which may be useful in  
 CC the treatment of disorders associated with the proteins of the  
 CC invention e.g., cancers and immunological disorders. The present  
 CC sequence represents cDNA encoding human TANGO protein of the invention.  
 CC  
 XX  
 S0 Sequence 813 BP; 164 A; 213 C; 235 G; 201 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1,26e-155 Length: 813  
 Score: 1446.00 Matches: 269  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.63% Mismatches: 0  
 Query Match: 99.79% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-905-674-2 (1-270) x AAF90657 (1-813)  
 QY 1 MethIstYrTyrArgTyrSerAsnAlaLysValSerCysTrrPyrLysTyrLeuLeuPhe 20  
 Db 1 ATGCACATATATATATATCTTAAGCCCAAGGTACAGCTGTGTGATCAAAATCTCTTTAC 60  
 QY 21 SerTyrAsnIleIlePheThrPleuAlaGIyValValPheLeuGIyValGIyLeuTrrPala 40  
 Db 61 AGCTACAAACATCATCTCTGCTGTGCTGAGATTGCTCTTCCTTGAGAGCGGGCTGGGCA 120  
 QY 41 TrrPserGIuLysGIyValLeuSerAspLeuThrLyValThrArgMetHisGIyIleAsp 60  
 Db 121 TGGAGCGAAAGGGGTGTGCTGTCCGACCTCAACAAATGACCCGGAGCTTGAATTCGAC 180  
 QY 61 ProValValLeuValLeuMetValGIyValValMetPheThrLeuGIyPheAgiLys 80  
 Db 181 CCTGTGTGCTGTGCTGTGATGGGCGGTGATGTTCACCTGGGGTTCGCCGGCTGC 240  
 QY 81 ValGIyAlaLeuArgLysnIleCysLeuLeuAsnPhenPheCysGIyThrIleValLeu 100  
 Db 241 GTGGGGGCTCTGGGAGATATCTGCTGCTCAACTTTTGTGTGGCCATCTGCTGC 300  
 QY 101 IlePhePheLeuGIuLeuAlaValAlaValLeuAlaPheLeuPheGlnAspIrrValArg 120  
 Db 301 ATCTTCTCTGAGCTGTGCTGTGGCCGTGCTGGCTTCTTCTTCAGAGCATGGGTGAGG 360  
 QY 121 AspArgPheArgLysPhePheGlnSerAsnIleLysSerTyrArgAspAspIleAspLeu 140



Alignment Scores:			
Pred. No.:	1,64e-155	Length:	813
Score:	1445.00	Matches:	269
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.63%	Mismatches:	0
Query Match:	99.72%	Indels:	0
DB:	22	Gaps:	0
US-09-905-674-2 (1-270) x AAF90659 (1-813)			
QY	1 MethiSTYrTYrArGTySeraSnAlaYsValSerCysTriPTyrlYstYrLeuLeuphe	20	
Db	1 ATGCACATATTATNAGATACCTCAACGCCAAGGTCAGCTGCTGTACAGACTCTTTTC	60	
QY	21 SerTYrAsnLlellePheTrLeuAlaGlyValAlaPheLeuGlyValAlcYLeuTrpAla	40	
Db	61 AGCTACAAATCATCTTCTGTGGCTGGAGTGTCTTCTTGAGATGGGCTGTGGCA	12	
QY	41 TrpSerGluLysGlyValLeuSerAspLeuTrpYsValTThrArgMetHisGlyLLeasp	60	
Db	121 TGGAGCGAAAGGGTGTGCTGTGCCAGCTCACCAAGATGACCAGATGATGAATCGAG	18	
QY	61 ProValAlaLeuValLeuMetValGlyValAlaMetPheTrLeuGlyPheAlaGlyCys	80	
Db	181 CCTGTGTCTGTCTGTCTGTGTGGCGCTGTGATGTTTCACCTGGGGTGTGGCGGTGC	24	
QY	81 ValGlyAlaLeuArgGluAsnLleCysLeuLeuAsnPhePheCysGlyTThrLleValLeu	100	
Db	241 GTGGGGGCTCTCGGGAGATATATGCTTGCTGCACCTTTTCTGTGGACCACTCTGCTC	30	
QY	101 IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGluAspTrpValArg	120	
Db	301 ATCTTCTTCTGTGAGCTGTGCTGTGGCCGCTGTGGCTTCTGTGTCCAGAGACTGGGTAGG	36	
QY	121 AspArgPheArgGluPhePheGluSerAsnLleLysSerTYrArgAspAspLLeaspLeu	140	
Db	361 GACCGGTTCTCGGGAGATTTCTTCAGAGCAACATCAAGTCTTACCCGGAGATATGATCTG	420	
QY	141 GluAsnLeuLLeaspSerLeuGluLysAlaAsnGlnCysGlyValATyGlyPProGlu	160	
Db	421 CAAACCTCATGACTCCCTTACGAAAGCTAACACAGTCTGTGGGGCATATAGCCCTTGA	480	
QY	161 AspTrpAspLeuAsnValITyPheAsnCysSerGlyAlaSerTYrSerArgGluLysCys	180	
Db	481 GACTGGGAGCTTCAAGTGTACTTCAATTGACAGCGGTGCACCTACAGCCGAGAAAGTGC	540	
QY	181 GlyValIProPheSerCysCysValIProAspProAlaGlnLysValAlaAsnThnGlnCys	200	
Db	541 GGGGTCCTTCTCTGTGTGGCTGTGGCAATCTGTGGCAAAAAGTTGTGATACACAGTGT	600	
QY	201 GlyTYrAspValArgLleGluLeuLysSerLysTrpAspGluSerIlePheTrhLysGly	220	
Db	601 GGAATGTGATGTGCAGATTACACTGMAAGGCAAGTGGATGTGATCATCTTACAGAAAGGC	660	
QY	221 CysLleGlnAlaLeuGlnSerTrpLeuProArgAsnLleTrLleAlaAlaGlyValPhe	240	
Db	661 TGCATCCAGCGGCTGTGAAAGCTGGCTCCCGCGGACATTTTCAATTGTGGCTGTTC	720	
QY	241 IleAlaLleSerLeuLeuGlnIlePheGlyLlePheLeuAlaArgTrhLeuLleSerAsp	260	
Db	721 ATCGCATCTGTGCTGTGTGGCAATTTTGGCATCTTCTTCGCAAGAGAGCTGATCTCAGAC	780	
QY	261 IleGluAlaValLysAlaGlyHisHisPhe 270		
Db	781 ATCGAGCGAGTGAAGCGCGGCATCATCTTC 810		
RESULT 8			
AAZ65259 standard; DNA: 2672 BP.			
ID	AAZ65259		
XX	AAZ65259		
XX	AAZ65259		

Percent Similarity:	99.63%	Conservative:	0
Best Local Similarity:	99.63%	Mismatches:	1
Query Match:	99.10%	Indels:	0
DB:	21	Gaps:	0
US-09-905-674-2 (1-270) x AA65259 (1-2672)			
QY 1 MethistyrinargyrsersasAlaIalysValSerCysrtrpIlyrIlyrLeuLeuPhe	20		
Db 178 ATGACATATATATATATCTCTAACGCCAAGTCAAGCTGCTGATACAAATACCTCTTTTC	237		
QY 21 SerIyrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAla	40		
Db 238 AGCTAACAAATCATCTCTGTTGGCTGGAGTTGTCTCTTGGAGTCCGGCTGTGGCA	297		
QY 41 TrpSerGlyIysGlyValLeuSerAspLeuThrIlyValThrtrgmethIasGlyIleAsp	60		
Db 298 TGGAGCCGAAAGGGGTGTCTGTCCGACCTACCAAAATGACCCGGATGCTGGAAATCGAC	357		
QY 61 ProValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys	80		
Db 358 CCTGTGGTGTGCTGTCTGTATGGTGGGGGTGTATGTTCACCCGGGGGTTCGGCGCTGC	417		
QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu	100		
Db 418 GTGGGGGCTGTCCGGGAGAAATATCTGCTCAACTTTTCGTGTGCACATCGTGC	477		
QY 101 IlePhePheLeuGluIleuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg	120		
Db 478 AATCTTCTTCGTGGAGCTGGCTGTGGCTGTGGCTCTTCTCTTCCAGAGCTGGGTAGC	537		
QY 121 AsparGpheaTrgGluPhePheGluSerAsnIleIlysserIyrArgAspIleAspLeu	140		
Db 538 GACCGGTTCCGGGAGTCTTCTGAGAGCAATCAAGTCTTACCGCGAGATATCGATCTG	597		
QY 141 GlnAsnLeuIleAspSerLeuGlnIlyAlaAsnGlnCysGlyAlaTrgIlyProGlu	160		
Db 598 CAAAACCTCATCGACTCCCTTCAGAAACATACCAAGCTGTGTGGCATATGGCCCTGAA	657		
QY 161 AspTrpAspLeuAsnValIyrPheAsnCysSerGlyAlaSerIyrSerArgGluIlyCys	180		
Db 658 GACTGGGACTTCACGCTTACTTCAATTGCAAGGGGTGCAGCTACGCCGAGAAATGTC	717		
QY 181 GlyValProPheSerCysCysValProAspProAlaGlnIlyValAlaAsnThrGlnCys	200		
Db 718 GGGGTCCTCTCTCTCTGCTGCGTGCACAAATCTCTCGCAAAAAGTGTGACACACACATGT	777		
QY 201 GlyIyrAspValArgIleGlnLeuIlySerIyrTrpAspGluSerIlePheThrIlyGly	220		
Db 778 GGATATGATGTAGGATTCAGCTGAAAGCAAGTGGGATGGTGCATCTTCACCAAAAGCC	837		
QY 221 CysIleGlnAlaIleGluSerTrpLeuProArgAsnIleIyrIleValAlaGlyValPhe	240		
Db 838 TGCATCCAGCGGCTGGAAAGTGGCTCCCGGGAACATTTACATTTGGCGTGGCTTTC	897		
QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrIleIleSerAsp	260		
Db 898 ATCGCATCTCCCTTGCAGATATTGGCATCTTCTCGCAAGACCGCTGATCTCAGAC	957		
QY 261 IleGluAlaValIlyAlaGlyHisHisPhe	270		
Db 958 ATCGAGCGCATGAAGCCGCGCATCACTTC	987		
RESULT 9			
AA592016			
ID AA592016 standard; cDNA; 2623 BP.			
AC AA592016;			
XX			

XX	Human: chromosome mapping; gene mapping; gene therapy; forensic:
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
OS	Homo sapiens.
PN	WO200175067-A2.
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	(HYSE-) HYSEQ INC.
PA	
PI	Dermanac RT, Liu C, Tang YT;
XX	
DR	WPI: 2001-639362/73.
DR	P-PDB: ABG27829.
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
XX	
PS	Claim 1; SEQ ID No 27820; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human
CC	diagnostic coding sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 2623 BP; 515 A; 713 C; 719 G; 659 T; 17 other;
	Alignment Scores:
	Pred. No.:            1.72e-152            Length:         2623
	Score:                1425.00            Matches:          267
	Percent Similarity:   98.89%           Conservative:      3
	Best Local Similarity:   98.89%           Mismatches:       0
	Query Match:           98.34%           Indels:            0
	DB:                    23                Gaps:              0
US-09-905-674-2 (1-270) x AMS92016 (1-2623)	
OY	1 Methistyrtyrtyrtyrserrsnalalyvalsercytprtyrlvstyrlleuleuphe 20
Db	67 ATCGGTGAATGAATGACTTCACGCCAGAGTGGTGCTGATCAACAATCCTCTTTTC 126
OY	21 SerTyAsnAlleIlelphetrPLeuAlaGlYValVaIPheLeuGlYValGlYLLeuTPala 40
Db	127 AGCTAACACATCATCTCTCGTGGTGGCGAGATTGTTCTTCTTGAGAGTGGGGCTGTGGGC 186
OY	41 TrpSerGIuLysGlyValLeuSerAspLeuThrLYsValThraArgMetHisGlyIleasp 60
Db	187 TGAGGCGAAAGAGGGTGTCTGTCCGACCTTCACCAAAGTAGACCAGGATGCAATGAGATGAC 246

OY	61	ProValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys	80
Db	247	CCGTGTGGTGGTGGTCCGTGATGGTGGGCGTGATGTTCACCTCTGGGTGGCGGGCTGC	306
OY	81	ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu	100
Db	307	GTGGGGGGCTCTGGCGGAGAAATATCTGCTTCTCAACTTTTCTGTGGACCAATCTGTGCTC	366
OY	101	IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg	120
Db	367	ATCTTCTCTCCGGAGCGGTGGCTGTGGCCGAGCTGGCCCTTCTCTTCAGAGACTGGGTGAGG	426
OY	121	AspTrpPheArgGluPhePheGlnSerAsnIleLeuSerTrpArgAspAspIleAspLeu	140
Db	427	GACCGTTCCGGGAGCTTCTTCGAGAGCAATCAATCTCTACCGGGAGCATATCATCTGTG	486
OY	141	GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGlu	160
Db	487	CAAAACCTCATCGCTCTCCCTTCAGAAAGCTAACCACTGCTGTGGGCAATATGGCCCTTAA	546
OY	161	AspTrpAspLeuAsnValTrpPheAsnCysSerGlyAlaSerTrpSerArgGluLysCys	180
Db	547	GACTGGGAGCTCAACAGCTACTTCAATTGTGAGGGGTGGCCAGCTACAGCCGAGAGAAAGTC	606
OY	181	GlyValProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCys	200
Db	607	GGGGTCCCTTCTCTCTCTGCTGCGCCAGATCTCGCCCAAAAGTTGTGCACACACAGTGT	666
OY	201	GlyTrpAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly	220
Db	667	GGATRTGATGTGCAGATTACGCTGACGTCGAGACCACTGGATGATCTCATCTTCCAGAAAGC	726
OY	221	CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTrpIleValAlaGlyValPhe	240
Db	727	TGCATTCAGAGCGGCTGGAAAGCTGGCTCCCGCGGAAATTTACATTGTGGCTGGCTCTTC	786
OY	241	IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp	260
Db	787	ATCGGCATCTCGCTGTGGTGCAGATATTTGGCATCTTCTCTGGCAAGACGCTGATCTCAAC	846
OY	261	IleGluAlaValLysAlaGlyHisHisPhe	270
Db	847	ATCAGAGCAGTGAAGCGCGCCATCACTTC	876
RESULT 10			
AAA93623			
ID	AAA93623 standard; DNA; 1428 bp.		
XX	AAA93623:		
AC			
XX			
DT	16-JAN-2001 (first entry)		
XX			
DE	Human tetraspanin-like protein splice variant SECX 3911675 DNA.		
XX			
KW	SECX protein: human; secreted; membrane-associated; cancer;		
KW	proliferation regulator; differentiation regulator; non-malignant tumour;		
KW	immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;		
KW	infection; inflammatory disorder; arthritis; hemopoietic disorder;		
KW	skin disorder; cardiovascular disorder; atherosclerosis; restenosis;		
KW	neurological disease; Alzheimer's disease; trauma; wounding;		
KW	spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;		
KW	anti-HIV; antiinflammatory; antiarthritis; antilariclosclerotic;		
KW	neuroprotective; vulnerable; antiatheritic; antimicrobial; cardiant;		
KW	dermatological; gene therapy; ds.		
XX			
OS	Homo sapiens.		
XX	—		
PN	WO200053742-A2.		
XX			
PD	14-SEP-2000.		
XX			
PF	09-MAR-2000; 2000WO-0506280.		

XX 09-MAR-1999; 99US-0123667.  
 PR 08-MAR-2000; 2000US-0123667.  
 XX (CURA-) CURAGEN CORP.  
 PA  
 PI Shimkets RA;  
 XX WPI: 2000-594318/56.  
 DR P-PSDB: AAB23036.  
 XX  
 PT Novel human membrane associated or secreted polypeptides and  
 PT polynucleotides useful for diagnosis, prevention and treatment of  
 PT pathological states such as cancer, immune, cardiovascular and  
 PT neurological disorders  
 XX  
 PS Claim 3; Fig 8A; 151pp; English.  
 XX Sequences AAA93616-A93631 and AAA93673-A93676 represent nucleic acids  
 CC which encode human SECX proteins (AAB23029-B23048). The SECX proteins  
 CC of the invention are either secreted or membrane-associated proteins  
 CC and act as regulator of cellular proliferation and differentiation. SECX  
 CC proteins or nucleotides are useful for diagnosing the presence of, or  
 CC predisposition to, a disease associated with altered levels of SECX  
 CC proteins and nucleotides. The SECX proteins are also useful to screen  
 CC compounds that modulate SECX activity or expression. The interaction of  
 CC a SECX protein with other cellular proteins may be useful to modulate  
 CC the activity of a partner protein, cellular proliferation, cellular  
 CC differentiation and cell survival. SECX nucleotides are useful for the  
 CC recombinant expression of SECX protein, and may be used to detect SECX mRNA  
 CC or genetic lesions in the SECX gene. They may also be used to modulate  
 CC SECX expression (e.g., using antisense oligonucleotides). SECX nucleic  
 CC acid sequences are also useful for identifying a cell or tissue type in  
 CC a biological sample, and in forensic biology. SECX primers or probes are  
 CC useful for detecting the presence of SECX nucleotides and for screening  
 CC tissue cultures for contamination. Diseases that may be treated or  
 CC prevented using SECX proteins or nucleotides include cancer (e.g.,  
 CC colorectal carcinoma, prostate cancer), benign tumours, immune disorders  
 CC (including autoimmune diseases, transplant rejection, allergies, AIDS),  
 CC infections, inflammatory disorders, arthritis, hematopoietic disorders,  
 CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis,  
 CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,  
 CC surgical or traumatic wounds, spinal cord injury), and skeletal  
 CC disorders.  
 XX  
 XX Sequence 1428 BP; 279 A; 408 C; 426 G; 315 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 3,47e-152 Length: 1428  
 Score: 1419.00 Matches: 265  
 Percent Similarity: 98.15% Conservative: 0  
 Best Local Similarity: 98.15% Mismatches: 5  
 Query Match: 97.93% Indels: 0  
 DB: 21 Gaps: 0  
 US-09-905-674-2 (1-270) x AAA93623 (1-1428)  
 QY 1 MethISTYTYRARGTYSERASNALYVALSERCYSTRPYLYLTYRLEUPE 20  
 DB 96 ATGCACTATTATGATACCTTAACGCCAAGGTGACCTCTGTTACAGTACCTCTTTTC 155  
 QY 21 SerTYAsnIleIlePheThrPLeuAlaGlyValPheLeuGlyValGlyLeuTrpAla 40  
 DB 156 AGCTACACATCATCTTCTTGCTGGTGGAGTGTCTTCCTTGAGGTGGGGTGGGCA 215  
 QY 41 TrpSerGluGlyGlyValLeuSerAspLeuThrIleValThrArgMetHisGlyIleAsp 60  
 DB 216 TGGAGCGAAAAGGGGTGCTGCCGACCTCACCAAGTGACCCGGATGCATGGAAATTCAC 275  
 QY 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80  
 DB 276 CATTGGCTGCTGCTCTGATGGTGGGCGCGGTGATGTTTACCCCTGGGGTTCGCCGCGCGC 335

QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhenPheCysGlyThrIleValLeu 100  
 DB 336 GTGGGGGCGCGGAGGAGATATCTGCTTGCATCACTTTTCTGTGGACCATCGCTGCTC 395  
 QY 101 IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg 120  
 DB 396 ATCTTCTTCTCGAGCTGCTGCTGGCCGTGCGCTTCCTGTTCCAGACTGGGTGAGG 455  
 QY 121 AspArgPheArgGluPhePheGluSerAsnIleCysSerTYArgAspAspIleAspLeu 140  
 DB 456 GACCGGTCCGGAGATTCCTTCAGAGCAACATCACTCTTACCGGAGACATATGATCATCTG 515  
 QY 141 GlnAsnLeuIleAspSerLeuGlnIlyAlaAsnGlnCysCysGlyAlaTYrGlyProGlu 160  
 DB 516 CAAACCTCATGACATCCCTTCAGAAAGCTAACACAGTCTGTGGCGCATATGGCCCTTAA 575  
 QY 161 AspTrpAspLeuAsnValTYrPheAsnCysSerGlyAlaSerTYSerArgGluIlyCys 180  
 DB 576 GACTGGGACCTCAACGCTCACTTCAATTCACGGGTGCTCCAGCTACAGCCGAGAGAAATGC 635  
 QY 181 GlyValProPheSerCysCysValProAspProAlaGlnIlyValValAsnThrGlnCys 200  
 DB 636 GGGGTCCCTTCTCTCTCTGCTGCTGCGCAGATCTCGCCCAAAAGTTGTGAAACACAGTGT 695  
 QY 201 GlyTYrAspValArgIleGlnLeuIlySerTYTrpAspGluSerIlePheThrIlyGly 220  
 DB 696 GGATATGATGTGACAGATTCAGTGAAGAGCAAGTGGATGAGTCCATCTTCACGAAAGGC 755  
 QY 221 CysIleGlnAlaLeuGluSerTYTrpLeuProArgAsnIleTYrIleValAlaGlyValPhe 240  
 DB 756 TGCATCCAGGGCGCTGGAACCTGGCTCCCGGGAACATTTTACATTGTGGCTGGCTTTC 815  
 QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgTYrLeuIleSerAsp 260  
 DB 816 ATCGCATCTGCTGCTTGCAGATATTTGGCATCTTCTCGCAAGCAGCTGATCTCAGAC 875  
 QY 261 IleGluAlaValIlyAlaGlyHisHisPhe 270  
 DB 876 ATCGAGCGACGTGAAGACCGGCCATCACTTC 905  
 RESULT 11  
 AA265350  
 ID AA265350 standard; DNA; 1667 BP.  
 XX  
 AC AA265350;  
 XX  
 DT 23-MAR-2000 (first entry)  
 XX  
 DE Human secreted protein gene 10 fragment.  
 XX  
 KW Human; secreted protein; cancer; tumour; developmental abnormality;  
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;  
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;  
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;  
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;  
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;  
 KW therapy; chromosome 10; ds.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 PN W09958660-A1.  
 PD 18-NOV-1999.  
 XX  
 PF 06-MAY-1999; 99WO-US09847.  
 XX  
 PR 12-MAY-1998; 98US-0085093.  
 PR 12-MAY-1998; 98US-0085094.  
 PR 12-MAY-1998; 98US-0085105.  
 PR 12-MAY-1998; 98US-0085180.  
 PR 18-MAY-1998; 98US-0085906.  
 PR 18-MAY-1998; 98US-0085920.  
 PR 18-MAY-1998; 98US-0085921.  
 PR 18-MAY-1998; 98US-0085921.

PR 18-MAY-1998; 98US-0085922.  
 PR 18-MAY-1998; 98US-0085923.  
 PR 18-MAY-1998; 98US-0085924.  
 PR 18-MAY-1998; 98US-0085928.  
 PR 18-MAY-1998; 98US-0085925.  
 PR 18-MAY-1998; 98US-0085927.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;  
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;  
 PI Latleir DW, Endress GA, Ebner R;  
 XX WPI, 2000-062296/05.  
 DR P-PSDB; AAY76266.  
 XX  
 PT New isolated human genes and the secreted polypeptides they encode,  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 XX  
 PS Disclosure: Page 427-428; 475pp; English.  
 XX  
 CC AA265250 to AA265350 represent 97 isolated human secreted protein genes.  
 CC This sequence was found to be present on human chromosome 10.  
 CC AAY76124 to AAY76223 represent the secreted proteins encoded by the 97  
 CC human genes. The genes and their corresponding secreted polypeptides are  
 CC useful for preventing, treating or ameliorating medical conditions,  
 CC e.g. by protein or gene therapy. Also pathological conditions can be  
 CC diagnosed by determining the amount of the new polypeptides in a sample  
 CC or by determining the presence of mutations in the new genes. Specific  
 CC uses are described for each of the 97 genes, based on which tissues they  
 CC are most highly expressed in, and include developing products for the  
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities  
 CC and foetal deficiencies, blood disorders, diseases of the immune system,  
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive  
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin  
 CC disorders, schizophrenia, diabetes, cardiovascular disorders, kidney  
 CC disorders, digestive/endocrine disorders, infections and AIDS. The  
 CC polypeptides are also useful for identifying their binding partners.  
 CC The sequences shown in AAY76224 to AAY76424 represent fragments of the  
 CC secreted proteins.  
 CC  
 XX  
 SQ Sequence 1667 BP; 323 A; 469 C; 488 G; 387 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 4.67e-147 Length: 1667  
 Score: 1375.00 Matches: 267  
 Percent Similarity: 98.53% Conservative: 1  
 Best Local Similarity: 98.16% Mismatches: 2  
 Query Match: 94.89% Indels: 3  
 DB: 21 Gaps: 0  
 US-09-905-674-2 (1-270) x AA265350 (1-1667)  
 QY 1 MethISTYTYRARGTYRSerASnAlaLysValSerCysTrpTyrIleuLeuPhe 20  
 DB 49 ATGCACATATTAATCTTAACGCGCAAGTCACTGCTGATCAAGTCACTCTTTTC 108  
 QY 21 SerTYRAsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAla 40  
 DB 109 ACCTACACATCACTCTCTGATGTGGCTGAGTTGCTCTTCTGAGTCCGGCGCTGGCGCA 168  
 QY 41 TrpSerGlyLysGlyValLeuSerAspLeuThrLysValThrTrpMetHisGlyIleAsp 60  
 DB 169 TGGACGCAAAAGAGTGTCTCTCCGACCTCAACCAAGTGAACCGGATGATGGAATCCGAC 228  
 QY 61 ProValIleuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys 80  
 DB 229 CCGTGGTGTGCTGTCTGATGTGGCTGATGTGGCTGATGTGGCTGATGTGGCTGATGTGG 288  
 QY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnIlePheCysGlyThrIleValLeu 100  
 DB 289 GTGGGGCTCTGCGGGAATATCTGCTGCTCAACTTTTGTGTGGCACCATCGTGTCTC 348

QY 101 IlePhePheLeuGlyLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg 120  
 DB 349 ATCTTCTTCTGAGACCTGCTGTGGCCCTGCTGCTCTCTCTGTTCCAGGACTGTGGTGAAG 408  
 QY 121 AspArgPheArgGlyPhePheGlyLeuSerAsnIleLysSerTYRArgAspAspIleAspLeu 140  
 DB 409 GACCGGTTCCGGGAGTCTCTGAGAGCAACATCAAGTCCATACCGGAGCATATCATCTCG 468  
 QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTrpGlyProGlu 160  
 DB 469 CAAMACCTATGACTCTCTCTGAGAACTCAACAGCTGTGGCCATATGCTGAA 528  
 QY 161 -AspTrpAspLeu-AsnValTYRPhaEncysSerGlyAlaSerTYRArgGlyLysC 180  
 DB 529 AGACTGGACCTCAGACGCTACTTCAATTGACAGGATGCCAGCTACAGCGGAGAA-T 587  
 QY 180 YSGIlyValProPheSerCysCysValProAspProAlaGlnLysValValAsnThrGlnC 200  
 DB 588 GCGGGGTCCCTTCTCTGCTGCTGCGCCAGATCCTGCGCAAAAGTTGTGAACACACAGT 647  
 QY 200 YSGIlyTyraAspValArgIleGlnLeuLysSerLysTrpAspGlySerIlePheThrLysG 220  
 DB 648 GTGATATATATCTCAGATTCAGCTGAGAGACAGTGAAGTCAATCTTTCACCAAG 707  
 QY 220 LysCysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTYRTrpIleValAlaGlyValP 240  
 DB 708 GCTGCATCCAGCGCTGGAACCTGCGTCCCGGCAACATTTACATTGTGGCTGGCGTCT 767  
 QY 240 IleIleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerA 260  
 DB 768 TCAATGCCATTCGCTGCTGTTCGAGATATTTGGCATCTTCTGCGCAAGACCTATCTCAG 827  
 QY 260 sPTleGlnAlaValLysAlaGlyHisHisPhe 270  
 DB 828 ACATCGAGCAGTGAAGCGCGCCATCACTTC 859  
 RESULT 12  
 ID AAA60182 standard; DNA; 639 BP.  
 XX  
 AC AAA60182;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 DE Hydrophobic domain protein cDNA HP02956 isolated from KB cells.  
 XX  
 KW Human; secreted protein; membrane protein; hydrophobic domain;  
 KW proliferation control; differentiation induction; material transport;  
 KW biophysics; signal receptor; ion channel; transporter; immunostimulant;  
 KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;  
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200029448-A2.  
 XX  
 PD 25-MAY-2000.  
 XX  
 PE 17-NOV-1999; 99WO-JP06412.  
 XX  
 PR 17-NOV-1998; 98JP-0326255.  
 PR 22-DEC-1998; 98JP-0364315.  
 PR 16-MAR-1999; 99JP-0069811.  
 PR 27-APR-1999; 99JP-0119299.  
 PR 19-MAY-1999; 99JP-0138169.  
 XX  
 PA (SAGA) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Kimura T;



XX WPI: 2000-387753/33.  
 DR P-PSDB; AAB12120.  
 XX  
 PT Proteins comprising hydrophobic regions, such as secretory and membrane  
 PT proteins, useful in research and diagnostics and having various  
 PT activities e.g. immunomodulatory, antiinflammatory, chemokineic,  
 PT hemostatic, thrombolytic -  
 XX  
 PS Claim 3; Page 195; 410pp; English.  
 XX  
 CC Secretory proteins play important roles in the proliferation control, the  
 CC differentiation induction, the material transport and the biophylaxis of  
 CC cells. Membrane proteins have important roles as signal receptors, ion  
 CC channels and transporters. The present sequence is the coding sequence  
 CC for a human protein which has at least one hydrophobic domain. The  
 CC protein encoded by the present sequence may be a secretory or a membrane  
 CC protein. The encoded protein may have cytokine and cell  
 CC proliferation/differentiation activity, immune stimulating or suppressing  
 CC activity, haematopoiesis activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, anti-inflammatory activity and tumour  
 CC inhibition activity. The present sequence could therefore be used for  
 CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
 CC disease, and cancer via gene therapy.  
 XX  
 SQ Sequence 639 BP; 131 A; 163 C; 188 G; 157 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 1,5e-118 Length: 639  
 Score: 1120.50 Matches: 213  
 Percent Similarity: 78.89% Conservative: 0  
 Best Local Similarity: 78.89% Mismatches: 57  
 Query Match: 77.33% Indels: 1  
 DB: 21 Gaps: 1  
 US-09-905-674-2 (1-270) x AAA60182 (1-639)  
 QY 1 MethIstYrTyrArGTYrSerAsnAlaLysValSerCysTrpTYrLysTyrLeuPhe 20  
 Db 1 ATGCACATATATAGTACCTAACGCCAAGTCACTGTGTACAACTCTCTTTTC 60  
 QY 21 SerTYrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40  
 Db 61 AGCTACAAACATCATCTTGTGGTGGCTGAGTGTCTCTTGAGTGGGCTGTGGCA 120  
 QY 41 TrpSerGluysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60  
 Db 121 TGGAGCGAAAGAGGTGTGTCTGCCACTCACCAGAGTACCGGATGCATGCAATGCAC 180  
 QY 61 ProValValIleuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80  
 Db 181 CCGTGTGTCTGCTCTGATGTGGGCGTGGATGTTCCCTGGGTTGCCCGCTGTC 240  
 QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100  
 Db 241 GTGGGGGCTCTGGGAGAGATATCTGCTTCAACTTT----- 279  
 QY 101 IlePhePheLeuGluLeuAlaValAlaValIleuAlaPheLeuPheGluAspTrpValArg 120  
 Db 279 ----- 279  
 QY 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTYrArgAspAspIleAspLeu 140  
 Db 279 ----- 279  
 QY 141 GluAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTYrGlyProGlu 160  
 Db 280 -----AACCAAGTCTGTGGCGCATATAGCCCTGAA 309  
 QY 161 AspTrpAspLeuAsnValTYrPheAsnCysSerGlyAlaSerTYrSerArgGluLysCys 180  
 Db 310 GACTGGAGCTCAACGCTTACTTCAATTGCACAGGGTGCAGCTACAGCGAGAGAGTGC 369

QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValIleAsnThrGlnCys 200  
 Db 370 GGGGTCCCTTCTCTCGCTGCGCTGCCAGATCTCGCGCAAAAAGTTGCAACACACTGT 429  
 QY 201 GlyTYrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220  
 Db 430 GGATATGATGTCAGATTCAGCTGCAAGAGCAAGTGGATGAGTCCATCTTCACGAAGGC 489  
 QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTYrIleValAlaGlyValPhe 240  
 Db 490 TGCATCCAGCGCTGCAAGCTGCGTCCCGCGAACATTTACATTGTGCGTGGCTTTC 549  
 QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260  
 Db 550 ATGCCCATCTCCCTGTTCAGATATTTCGATCTTCTGCGCAAGAGAGCTGATCTCAGAC 609  
 QY 261 IleGluAlaValLysAlaGlyHisHisPhe 270  
 Db 610 ATGAGGCAGTGAAGCGCGCATCATCTTC 639  
 RESULT 13  
 ID AAA60192  
 XX AAA60192 standard; DNA; 2367 BP.  
 AC AAA60192;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 DE Hydrophobic domain protein cDNA HP02956 isolated from KB cells.  
 XX  
 KW Human; secreted protein; membrane protein; hydrophobic domain;  
 KW proliferation control; differentiation induction; material transport;  
 KW biophylaxis; signal receptor; ion channel; transporter; immunostimulant;  
 KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokineic;  
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200029448-A2.  
 XX  
 PD 25-MAY-2000.  
 XX  
 PF 17-NOV-1999; 99WO-JP06412.  
 XX  
 PR 17-NOV-1998; 98JP-0326255.  
 PR 22-DEC-1998; 98JP-0364315.  
 PR 16-MAR-1999; 99JP-0069811.  
 PR 27-APR-1999; 99JP-0119299.  
 PR 19-MAY-1999; 99JP-0138169.  
 XX  
 PA (SAGA ) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Kimura T;  
 XX  
 DR WPI: 2000-387753/33.  
 DR P-PSDB; AAB12120.  
 XX  
 PT Proteins comprising hydrophobic regions, such as secretory and membrane  
 PT proteins, useful in research and diagnostics and having various  
 PT activities e.g. immunomodulatory, antiinflammatory, chemokineic,  
 PT hemostatic, thrombolytic -  
 XX  
 PS Claim 4; Page 209-211; 410pp; English.  
 XX  
 CC Secretory proteins play important roles in the proliferation control, the  
 CC differentiation induction, the material transport and the biophylaxis of  
 CC cells. Membrane proteins have important roles as signal receptors, ion  
 CC channels and transporters. The present sequence is the coding sequence  
 CC for a human protein which has at least one hydrophobic domain. The



CC protein encoded by the present sequence may be a secretory or a membrane protein. The encoded protein may have cytokine and cell proliferation/differentiation activity, immune stimulating or suppressing activity, haematopoiesis activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, anti-inflammatory activity and tumour inhibition activity. The present sequence could therefore be used for treatment of autoimmune disease, Alzheimer's disease, Parkinson's disease, and cancer via gene therapy.

XX Sequence 2367 BP; 459 A; 673 C; 634 G; 601 T; 0 other;

#### Alignment Scores:

Pred. No.: 9.87e-118 Length: 2367  
Score: 1120.50 Matches: 213  
Percent Similarity: 78.89% Conservative: 0  
Best Local Similarity: 78.89% Mismatches: 0  
Query Match: 77.33% Indels: 57  
DB: 21 Gaps: 1

US-09-905-674-2 (1-270) x AAA60192 (1-2367)

OY 1 MethISTYRYRARGTYRSERASNAALAYSVALSERCYSTPYRYLYSTYRLEUPHE 20  
DB 69 ATGCACATATATAGTACTCTACGCGCAAGGTCAAGCTGCTGTACCACTCTCTTTC 128  
OY 21 SERTYRASNILLEPHEPTRLEUALAGLYVALVAPHELEUGLYVALGLYLEUTRALA 40  
DB 129 AGCTACAAACATCATCTTGTGCTGGCTGGAGTGTCTTCTTGAGAGTCGGGCTGGGCA 188  
OY 41 TRPSERGLULYSGLYVALLEUSERSASPLEUTHRLYSVALTHRARGMETHISGLYLLEASP 60  
DB 189 TGCAGCGAAAAGGGCTGCTGCTCCGACCTCACAAAGTGACCCGAGATGCATGAAATGCAC 248  
OY 61 PROVALLEUVALLEUMETVALGLYVALVAMETPHEPTRLEUGLYPHEALIGLYCS 80  
DB 249 CCTGTGGTGGTGGCTGATGATGCTGGCTGATGTTACACCGGGGTTGCGCGGCTGC 308  
OY 81 VALGLYALALEUARGGLUASNILECYSLEUENASNPHPECYSGLYTHRLILEVALLEU 100  
DB 309 GTGGGGGCTTCGGGAGAAATATCTGCTCTCAACTTT 347  
OY 101 ILEPHEPHEUGLULEUALVALAVALLEUALPHELEUPHEGLINASPTRVALARG 120  
DB 347 347  
OY 121 ASPARGPHEARGGLUPHEPHEGLUSERRASNILEYSSERTYRARGASPSPILEASPLEU 140  
DB 347 347  
OY 141 GLNASNLEULEASPSERLEUGLNLYSALASNGINCYSGLYVALIATYRGLYPROGLU 160  
DB 348 348  
OY 161 ASPTRPASPLEUASNVALTYRPHASNCYSSERCGLYALASERTYRSERARGGLULYS 180  
DB 378 GACTGGGACCTCAACGCTACTCTCAATTCAGCGGTGGCACCTACACCGAGAAATGTC 437  
OY 181 GLYVALPHEPHESERCYSCYSVALPROASPPROALAGLNLYSVALVALASNTHTGLN 200  
DB 438 GGGGTCCTTCTCTGCTGGTGGCTCCGAGATCTCTCGGAAAAAGTTGTGAACACACACAGTGT 497  
OY 201 GLYTRASPVALARGILEGLNLLEULYSERTYRTPASPGIUSERITLPEPHEPTRHLYSGLY 220  
DB 498 GGATATGATGTCAGGATTCACTAAGACGCAAGTGGGATGATGCTCATCTTCACGAAAGGC 557  
OY 221 CYSILLEGINALALEUGLUSERTYRLEUPROARGASNILETYRILEVALAAGLYVALPHE 240  
DB 558 TGCATCCAGCGCGCTGGAAGCTGGCTCCCGGGAACATTTCATTGGGCTGGGCTTTC 617  
OY 241 ILEALILESERLEUENGLNILEPHEGLYILEPHELEUALAAGTHRLLEULILESERASP 260  
DB 618 ATTCGCATCTCGTGTGTCACATATATTGGCATTTCTTGGCAAGGACGTGATCTCAGAC 677

OY 261 ILEGLUALVALYLSALAGLYHSHISHPHE 270  
DB 678 ATGCAGGCACTGAAGCGCGGCATCACTTTC 707

#### RESULT 14

AAA93675 standard; DNA; 816 BP.

AAA93675.

16-JAN-2001 (first entry)

Human tetraspantin-like protein splice variant SECX 3911675-2 DNA.

SECX protein; human; secreted; membrane-associated; cancer;  
proliferation regulator; differentiation regulator; non-malignant tumour;  
immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;  
infection; inflammatory disorder; arthritis; haematopoietic disorder;  
skin disorder; cardiovascular disorder; atherosclerosis; restenosis;  
neurological disease; Alzheimer's disease; trauma; wounding;  
spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
anti-HIV; antiinflammatory; antiarthritic; antiarteriosclerotic;  
neuroprotective; vulnerrary; antiallergic; antimicrobial; cardiant;  
dermatological; gene therapy; ds.

Homo sapiens.

WO200053742-A2.

14-SEP-2000.

09-MAR-2000; 2000MO-US06280.

09-MAR-1999; 9905-0123667.

08-MAR-2000; 200005-0123667.

(CURA-) CURAGEN CORP.

Shinkets RA.

WPI; 2000-594318/56.

P-PSTDB; AAB23047.

Novel human membrane associated or secreted polypeptides and polynucleotides useful for diagnosis, prevention and treatment of pathological states such as cancer, immune, cardiovascular and neurological disorders

Claim 3; Fig 8B; 151pp; English.

Sequences AAA93616-A93631 and AAA93673-A93676 represent nucleic acids which encode human SECX proteins (AAB23029-B23048). The SECX proteins of the invention are either secreted or membrane-associated proteins and act as regulator of cellular proliferation and differentiation. SECX proteins or nucleotides are useful for diagnosing the presence of, or predisposition to, a disease associated with altered levels of SECX proteins and nucleotides. The SECX proteins are also useful to screen compounds that modulate SECX activity or expression. The interaction of a SECX protein with other cellular proteins may be useful to modulate the activity of a partner protein, cellular proliferation, cellular differentiation and cell survival. SECX nucleotides are useful for the recombinant expression of SECX protein, and may be used to detect SECX mRNA or genetic lesions in the SECX gene. They may also be used to modulate SECX expression (e.g., using antisense oligonucleotides). SECX nucleic acid sequences are also useful for identifying a cell or tissue type in a biological sample, and in forensic biology. SECX primers or probes are useful for detecting the presence of SECX nucleotides and for screening tissue cultures for contamination. Diseases that may be treated or prevented using SECX proteins or nucleotides include cancer (e.g., colorectal carcinoma, prostate cancer), benign tumours, immune disorders (including autoimmune diseases, transplant rejection, allergies, AIDS), infections, inflammatory disorders, arthritis, haematopoietic disorders,

CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis  
CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,  
CC surgical or traumatic wounds, spinal cord injury), and skeletal  
CC disorders.

SQ Sequence 816 BP; 159 A; 230 C; 243 G; 184 T; 0 other;

**Alignment Scores:**

Pred. No.:	3.2e-110	length:	816
Score:	1049.00	Matches:	204
Percent Similarity:	97.61%	Conservative:	0
Best local Similarity:	97.61%	Mismatches:	4
Query Match:	72.39%	Indels:	2
DB:	21	Gaps:	0

US-09-905-674-2 (1-270) x AAA93675 (1-816)

QY	63	ValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCysValGly	82
Db	5	GTCCTGATCTCTGATGTGGGGCCGGGTGATGTCACCCG-GGGTTCCCGGGCCCGCTGGGG	63
QY	83	Ala-LeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhe	101
Db	64	GCGGGCCAGGGAAGATATCTGCTGCTCACTTTTCTGTGGCACCATGTGCTCATCTT	122
QY	102	ePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArgAspArg	122
Db	124	CTTCCGGAGAGCGCTGTGGCCGTGGCTGGCTTCTGTTCCAGAGCTGGGTAGGAGACCG	18
QY	122	gPheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGlnAs	14
Db	184	GTTCCGGAGATTTCTGAGAGCAATCAAGTCTTACCGGGAGCATATGATCTGCCAAA	24
QY	142	IleuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGluAspTr	16
Db	244	CCATCATCGACTCCCTTCAGAAAGCTAACCAAGTCTGTGGCGCATATGGCCCTGAAGACG	30
QY	162	PAAPLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCysGlyVal	18
Db	304	GGAACCTCAACGCTTACTTCAATTGAGGGGGCCAGCATACAGCCGAGAAAGCGGGGT	36
QY	182	IleProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCysGlyTyr	20
Db	364	CCCCCTCTCTGCTGGTGGTCCAGATCCCTGGCCAAAAGTTGTGACACACAGTGTGATA	42
QY	202	rasPvalArgIleGlnIleuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIle	22
Db	424	TGATGTCCAGGATTCACACTGAAAGCAAGCAATGGAGTCACTTCATTCACAAAAGCTGCAT	48
QY	222	eGlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAl	24
Db	484	CCAGGGCGCTGGAAAGCTGGCTCCCGGGACAACTTTACATTTGTGGCTGGCGTCTTCATCC	54
QY	242	AlaIleSerLeuLeuGlnIlePheGlyIlePheIleAlaArgThrLeuIleSerAspIleGln	26
Db	544	CATCTCGCTGGTGGACAGATATTGGCATCTTCTCGCAAGACCCGTATCTCAGACATCGA	60
QY	262	uAlaValLysAlaGlyIleHisIlePhe	270
Db	604	GCGAGTGAAGGCCCGGCATCATCTTC	628

RESULT 15

AC AAS80667;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16471.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss

XX	Homo sapiens.
OS	WO200175067-A2.
XX	
PN	11-OCT-2001.
XX	
PD	30-MAR-2001; 2001MO-US08631.
XX	
PF	31-MAR-2000; 2000US-0540217.
XX	
PR	23-AUG-2000; 2000US-0649167.
XX	
PPA	(HYSE-) HXSEQ INC.
XX	
PL	Dmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
XX	P-PSDB; ABG16480.
PT	New isolated polynucleotide and encoded polypeptides, useful in
XX	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
XX	biodiversity -
PS	Claim 1; SEQ ID No 16471; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human
CC	diagnostic coding sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 601 BP; 107 A; 168 C; 180 G; 146 T; 0 other;
Alignment Scores:	
Pred. No.:	1,72e-97 Length: 601
Score:	936.00 Matches: 175
Percent Similarity:	98.31% Conservative: 0
Best Local Similarity:	98.31% Mismatches: 3
Query Match:	64.60% Indels: 0
DB:	23 Gaps: 0
US-09-905-674-2 (1-270) x AAS80667 (1-601)	
QY	1 Methistyrtyrragtyrserasahalyvalsercystrrpyrltysryleuapne 20 
DB	67 AGCGCCTGTAGTACTACCTAACGCCAAGGTGGTGCGTAGCAAGTACCCTTTTC 126 
QY	21 SerTyraNiilellephetrpleualaglValyAlpheueuglyalglyleutrpala 40 
DB	127 ACCTACAACATCATCTTCTGCTTGGCTGGAGTTCTCTTCTTGGAGTGGGCTGTGGCA 186 
QY	41 TpsrsgiuLysgllyValleuSerAspLeuThrlsyalThrArpmethisilyleasp 60 
DB	187 TGGAACGAAAGGGGTGCTCTGCAGACCTCACCAAATGACCCGGATGATGAATCGAC 246 
QY	61 ProValValleuValleuMetValGlylValylValmetPheThrlleuglyPhealaglyCys 80 

Db	247	CGTGTGTCTGTCTCTTATGTGTGGGTGTGTATTTTACACCTTGGGGTTCGCGGCTGC	306
Oy	81	ValGlyAlaLeuArgLysAsnIleCysLeuLeuAsnPhenPheCysGlyThrIleValLeu	100
Db	307	GTTGGGGCTCTCGGGGAAATATCGTTGCTTCACATTTTCTGTGGCACAATCGTGCCTC	366
Oy	101	IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTyrValArg	120
Db	367	ATCTCTCTCTGTGACCTGAGCTGTGGCCGACCTGGCTTCTCTTCCAGGACTGGGTAGG	426
Oy	121	AspArgPheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeu	140
Db	427	GACCGTTCGGGAGTCTTCTTGAGAGCAACATCAAGTCTTACCGGGAGGATATGATCTG	486
Oy	141	GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyPProGlu	160
Db	487	CAAAACCTCATGACTCCCTTCAGAAAGCTTAACACAGTGTGGGGCATATGGCCCTGAA	546
Oy	161	AspTyrAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlu	178
Db	547	GACTGGGACCTACAGCTTACTTCAATTTCAGCGGTGGCCAGCTACAGCCGAGAG	600

Search completed: December 19, 2002, 06:26:57  
Job time : 269 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 19, 2002, 07:10:43 ; Search time 58 Seconds

(without alignments)  
1847.440 Million cell updates/sec

Title: US-09-905-674-2

Perfect score: 1449

Sequence: 1 MHYRYSNAKYSWCWKYLLF.....IFLARTLSDIEAVKAGHHF 270

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 356696 seqs, 198428768 residues

Total number of hits satisfying chosen parameters: 713392

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus\_model -DEV=x1h  
-Q=/cgn2\_1/USPTO\_spool/US0905674/runat.16122002.111847-6882/app\_query.fasta.1.455  
-DB=Published Applications NA -QPM=fastap -SUFFIX=tmph -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62  
-PPANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTEXT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US0905674.ecgn.1.1.21 @runat.16122002.111847.6882  
-NCPU=6 -ICPU=3 -NO\_XLPHY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGIOS  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PC7\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PC7US\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	867	59.8	1655	10 US-09-925-299-205
2	858	59.2	813	9 US-09-934-268-3
3	858	59.2	3184	9 US-09-934-268-1
4	841.5	58.1	2538	9 US-09-972-970-2

5	729.5	50.3	1110	10	US-09-729-674-169	Sequence 169, App
6	567	39.1	864	9	US-09-875-440-1	Sequence 1, Appl1
7	395.5	27.3	1778	12	US-10-052-886-277	Sequence 277, App
8	300.5	20.7	873	10	US-09-864-864-279	Sequence 279, App
9	294.5	20.3	1175	10	US-09-823-356-31	Sequence 31, Appl1
10	283	19.5	446	10	US-09-783-590-10353	Sequence 10353, A
11	282	19.4	690	10	US-09-895-828-430	Sequence 430, App
12	281	19.4	368	12	US-10-042-417-37	Sequence 37, Appl1
13	273.5	18.9	1797	10	US-09-729-674-89	Sequence 89, Appl1
14	273.5	18.9	2358	10	US-09-823-356-32	Sequence 32, Appl1
15	266	18.4	1545	10	US-09-925-301-392	Sequence 392, App
16	263.5	18.2	1725	10	US-09-822-830A-64	Sequence 64, Appl1
17	263.5	18.2	1792	9	US-09-954-531-1349	Sequence 1349, App
18	261.5	18.0	1204	10	US-09-822-849A-51	Sequence 51, App
19	259.5	17.9	2047	10	US-09-925-302-31	Sequence 31, Appl1
20	258.5	17.8	1125	10	US-09-967-768A-323	Sequence 323, App
21	257.5	17.8	1192	10	US-09-880-107-2351	Sequence 2351, App
22	257.5	17.8	1350	10	US-09-925-301-539	Sequence 539, App
23	257.5	17.8	1624	10	US-09-795-380-19	Sequence 19, Appl1
24	248.5	17.1	2018	10	US-09-822-830A-587	Sequence 587, App
25	245	16.9	732	9	US-10-025-380-1120	Sequence 1120, App
26	245	16.9	732	10	US-09-922-217-1120	Sequence 1119, App
27	245	16.9	1341	9	US-10-025-380-1119	Sequence 1119, App
28	245	16.9	1341	10	US-09-922-217-1119	Sequence 1052, App
29	243	16.8	1104	9	US-10-025-380-1052	Sequence 1052, App
30	243	16.8	1104	10	US-09-922-217-1052	Sequence 1052, App
31	243	16.8	1104	10	US-09-833-263-1052	Sequence 1052, App
32	243	16.8	1104	10	US-09-880-107-2343	Sequence 2343, App
33	243	16.8	1317	10	US-09-925-301-193	Sequence 193, App
34	237	16.4	1496	10	US-09-969-708-132	Sequence 132, App
35	235.5	16.3	1219	10	US-09-822-830A-421	Sequence 421, App
36	225.5	15.6	1529	10	US-09-799-777-79	Sequence 79, Appl1
37	221	15.3	572	10	US-09-998-598-563	Sequence 563, App
38	218.5	15.1	1289	9	US-09-232-880-111	Sequence 111, App
39	218.5	15.1	1289	9	US-10-012-896-111	Sequence 111, App
40	218.5	15.1	1289	10	US-09-759-143-111	Sequence 111, App
41	218.5	15.1	1289	10	US-09-780-666-111	Sequence 111, App
42	218.5	15.1	1289	10	US-09-030-606-111	Sequence 111, App
43	218.5	15.1	1289	10	US-09-822-827-111	Sequence 111, App
44	218.5	15.1	1289	10	US-09-115-453-111	Sequence 111, App
45	218.5	15.1	1324	9	US-09-981-876-82	Sequence 82, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-925-299-205  
Sequence 205, Application US/09925299  
Patent No. US20020055627A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925, 299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124, 270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 205  
LENGTH: 1655  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1548)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc.feature  
LOCATION: (1559)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc.feature

```

: LOCATION: (1564)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc:feature
: LOCATION: (1623)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc:feature
: LOCATION: (1643)
: OTHER INFORMATION: n equals a,t,g, or c
: US-09-925-299-205

```

Alignment Scores:	
Pred. No.:	1,68e-95
Score:	867.00
Percent Similarity:	76.25%
Best Local Similarity:	59.00%
Query Match:	59.83%
DB:	10
	Gaps:
	0
	Length:
	1655
	Matches:
	154
	Conservative:
	45
	Mismatches:
	62
	Indels:
	0
	Gaps:
	0

US-09-905-674-2 (1-270) x US-09-925-299-205 (1-1655)

Oy	6	TysetrAsnAlalysValSerCysTrpIlylYtyrLeuLeuPheSerTyrAsnIlelle	25
Db	292	TACAAAGGCTCCTCAAGTCACGTGGTCCATCAACAATCTCATATTGGGCTTAAGTGATA	351
Oy	26	PheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTPalaTrpSerGluYsgly	45
Db	352	TTTGGTGTGGGAATACATTTCTTGGAAATGGACGTGGGCAATGAAAAGA	411
Oy	46	ValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProValIleVal	65
Db	412	GTTCTGTGCCAACACTCTTCTCCATCACCGCATCTCCGGCGCTTGCACCAGTTGGCTCTTC	471
Oy	66	LeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArg	85
Db	472	CTTGCGTGGGAGGAGATGATGTCATTGGATTGGCATTTGCAGSGTGATTCGAGCGGTACGG	531
Oy	86	GluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGlu	105
Db	532	GA AAAA CACTTCTCTCCAGATT TTTTCTGCTGCTCCGGAAATTA TTCTCTCTGGAG	591
Oy	106	LeuAlaValAlaValIleuAlaPheLeuPheGlnAspTrpValArgAspArgPheArgGlu	125
Db	592	CTCACGTCCGGAGTTCAGATTTGGTTTTTCAAGACATGGATCAAGACCACTGTATTTC	651
Oy	126	PhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGlnAsnLeuIleAsp	145
Db	652	TTTATTAACACACATCATCAGACATATCGGGATGACATTTGTCAAAACCTCATAC	711
Oy	146	SerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGluAspTrpAspLeuAsn	165
Db	712	TTTCAACCCAGAAATTTGGCAGTGTGTGGGGCTTTTGGAGCGATGATGGAACTAAT	771
Oy	166	ValTyrPheAsnCysSergIylAlaSerTyrSerArgGluLysCysGlyValProPheSer	185
Db	772	ATTTCCTCATTTGCACAGATTCCAATGC AAGTCAGAGACGATGTCGGCTTCATTCTCC	831
Oy	186	CysCysValProAspProAlaGlnLysValAlaAsnThrGlnCysGlyTyrTrsPalaArg	205
Db	832	TGCTCTCACTAAAGATCCCGCGAAGATGTCATCAACATCAGTGGCGATTCATCCACGG	891
Oy	206	IleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGlnAlaLeu	225
Db	892	CAAAAACCAAGAGTTGCACACACAGATTGATTTACACGAAGGCTGTGTGCCCAATT	951
Oy	226	GluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeu	245
Db	952	GGAAGATGCTGCCAGACAATTTACCATCTGCTGTGATTTTCACTACGACATTGCATTCG	1011
Oy	246	IleGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlnAlaValLys	265
Db	1012	CTGCAGATATTGGGATATGCTGTGGCCCAAAATTTGGTAGCATATTCGAACCTGTACAG	1071
Oy	266	Ala 266	

```

      DB          1072 GCG 1074
      RESULT 2
      US-09-934-268-3
      Sequence 3, Application US/09934268
      Patent No. US20020172986A1
      GENERAL INFORMATION:
      APPLICANT: Leahy, Kevin R.
      TITLE OF INVENTION: 23228, A NOVEL HUMAN TETRASPANIN FAMILY
      TITLE OF INVENTION: MEMBER AND USES THEREOF
      FILE REFERENCE: 10448-079001
      CURRENT APPLICATION NUMBER: US/09/934,268
      CURRENT FILING DATE: 2001-08-21
      PRIOR APPLICATION NUMBER: 60/226,612
      PRIOR FILING DATE: 2000-08-21
      NUMBER OF SEQ ID NOS: 4
      SOFTWARE: FastSeq for Windows Version 4.0
      SEQ ID NO 3
      LENGTH: 813
      TYPE: DNA
      ORGANISM: Homo sapiens
      US-09-934-268-3

```

### Alignment Scores:

Pred. No.:	7.15e-95	Length:	813
Score:	858.00	Matches:	154
Percent Similarity:	75.76%	Conservative:	46
Best Local Similarity:	58.33%	Mismatches:	64
Query Match:	59.21%	Indels:	0
DB:	9	Gaps:	0

US-09-905-674-2 (1-270) x US-09-934-268-3 (1-813)

QY	3	TyrTyrArgTyrSerAsnAlaLysValSerCysTrpPyrItyrItyrLeuLeuPheSerTyr	22
Db	13	CCCCAGCATTTCCAGAACCTGAGGTGGCGTCGGGGAAATATCTTCGTGTGGCTTC	72
QY	23	AsnIleIlePheTrpLeuAlaGlyValValPheLeuAlaGlyLeuTrpAlaTrpSer	42
Db	73	AACTATGTCTTCTGGGTCTGGAGCCCTGTTCCTGGCTATGGGCCCTCTGGGCTGGGCT	132
QY	43	GluTyrGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProVal	62
Db	133	GAGAAAGGGGTCTCTCGAATCATCTCACCGCTGACAGATCTGGAGGGCCCTTGACCCGCTG	192
QY	63	ValLeuValLeuMetValGlyValMetPheThrLeuGlyPheAlaGlyCysValGly	82
Db	193	TGGCTGTGTGGTGGTGGAGGCGCTCATGGCGTCTGGGCTTGGCTGGCTGCAATTCGG	252
QY	83	AlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhe	102
Db	253	GCCCTCCGGGAAACAACCTCTCTGCTCAAGTTTTTCTCCGTTTCCTCGGTTCATCTTC	312
QY	103	PheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArgAspArg	122
Db	313	TTCCTGGAGCTGGCAACAGGATCCTGGCCTTGTCTTCAAGAGACTGGATTCCGAGACAG	372
QY	123	PheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGlnAsn	142
Db	373	CTCAACCTTTCATCAACACCAACCTCAAGGCTTACCGGGAGCGACATTGACTCCAAAC	432
QY	143	LeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGluAspTrp	162
Db	433	CTCATTTGACTTGGCTCAGGAATATCTGGCTTGTCTGTGGAGCCGACGCCCAATGACTGG	492
QY	163	AspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCysGlyVal	182
Db	493	AACCTCAATATCTACTCAACTGACACTGTGAACCCCAACCCGAGCGCTCGGGGGTG	552
QY	183	ProPheSerCysCysValProAspProAlaGlnLysValIleValAsnThrGlnCysGlyTyr	202
Db	553	CCCTTCCTCTGGCTGAGGGAGGAGGATGTCCTCAACACCCCACTGTGGCTAC	612

[illegible]

Qy	43	GlulysgylValleuSerAspLeuThrLysValThrArgmethHisgylIleasProval	62
Db	251	GAGAAAGGGGTCTTCCTCCAAACATCTCTGCGGTGACCGACATGGGAGCCCTGACCCCTGTG	310
Qy	63	ValleuValleuMeValGlyValValMePheThrLeuGlyPheAlaGlyCysValGly	82
Db	311	TGGCTGTTTGAAGTGTGGGAGCGGTCAATGTCCGTGCTGGGCTTGGCGGCTGCATGGG	370
Qy	83	AlaLeuArgGluAsnIleCysLeuIleuAsnPhePheCysgylThrIleValleuIlePhe	102
Db	371	GCTCTCGGAGAACACTTTCCTCTCAAGTTTTCACAGTTCCTTGGCCATCTTC	430
Qy	103	PheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnsptRpyAlaArgasParg	122
Db	431	TTCTCGAGCGTGGACACAGGATCTTGCCCTCTCATTCAAAGCACTGGATTCCAGACACG	490
Qy	123	PheArgGluPhePheGluSerAsnIleLysSerTyrTrpArgasPArgIleasPleuGlnAsn	142
Db	491	CTCATTTCTTCATTAAACAAACAGCTACAGGCCATATGGAGTGCATTGACATTCACCTCCAAAC	550
Qy	143	LeuIleasPserLeuGlnLysAlaAsnGlnCysCysgylAlaTyrGlyProGluAspTrp	162
Db	551	CTCATTTGACTTTGGCTCAGAAATATTTGGTCTGCGTGGAGACCCGAGGCGCTAATGACTGG	610
Qy	163	AspleuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCysgylVal	182
Db	611	AACCTCATATCTTTTTCACATGCACGTGACTTGACCCGAGCGGAGCGCTCGGGGTG	670
Qy	183	ProPheSerCysCysValProAspProAlaGlnLysValAlaAsnTrpGlnCysgylTyr	202
Db	671	CCCTTCTCTCTCTGTGTGCAGGGACCCCTGCCAT----GTCTCAACACCCAGTGTGGCTAT	726
Qy	203	AspValArgIleGlnLeuLysSerLysTyrAspIleuSerIlePheThrLysgylCysIle	222
Db	727	GATGTCCGGCTCAACACTGAGACTGGACAGAGGCGTCCATACACCAAAAGGCTGTGTG	786
Qy	223	GlnIleLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValAlaPheIleAla	242
Db	787	GGCCAGTTTGGAAGTGGCTGCCAGACACCACTGATGCTGTGGGTGGCTTTGGGGCC	846
Qy	243	IleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlu	262
Db	847	ATGCTCTCCCTCCAGATCTTTGGTATGTGCTGGCCCAAGAACTGTGATGATCAATCAAG	906
Qy	263	AlaValLysAla 266	
Db	907	GCACTGAAGGCC 918	
RESULT 5			
US-09-729-674-169			
Sequence 169, Application US/09729674			
Patent No. US2001003935A1			
GENERAL INFORMATION:			
APPLICANT: Jacobs, Kenneth			
APPLICANT: McCoy, John M.			
APPLICANT: LaValle, Edward R.			
APPLICANT: Collins-Racie, Lisa A.			
APPLICANT: Evans, Cheryl			
APPLICANT: Merberg, David			
APPLICANT: Treacy, Maurice			
APPLICANT: Agostino, Michael J.			
APPLICANT: Steindinger II, Robert J.			
APPLICANT: Spaulding, Vikki			
APPLICANT: Wong, Gordon G.			
APPLICANT: Clark, Hillary			
APPLICANT: Fechtel, Kim			
APPLICANT: Genetics Institute, Inc.			
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM			
FILE REFERENCE: 6055-64X			
CURRENT APPLICATION NUMBER: US/09/729,674			
CURRENT FILING DATE: 2000-12-04			
PRIOR APPLICATION NUMBER: 09/539,330			

```

; PRIOR FITTING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 169
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-169

Alignment Scores:
Pred. No.: 4,35e-79 Length: 1110
Score: 729.50 Matches: 140
Percent Similarity: 68.20% Conservative: 38
Best Local Similarity: 53.64% Mismatches: 62
Query Match: 50.35% Indels: 22
DB: 10 Gaps: 2

US-09-905-674-2 (1-270) x US-09-729-674-169 (1-1110)
QY 6 TTTTSeAsnAlaAlaValSerCysGlyTrpIlyrIysThrLeuLeuPheSerTyAsnIleIle 25
Db 191 TTTCAAGAGGCTCTGAAGTCAAGTTGGTCATCAAAATCTTCATATTGGCTTCATATGTATA 250
QY 26 PheTrpLeuAlaGlyValValPheLeuGlyAlaGlyLeuTrpAlaTrpSerGlyValGly 45
Db 251 TTTTGGTTTGGGAATAAATCTTCGTAATTTGGAGCTGGAGCTGGCATGGAAATGAAAAAGA 310
QY 46 ValLeuSerAspLeuThrLysValIthrArgmethIsgIlyLeaspProValValLeuVal 65
Db 311 GTTCTGTCCAAATCTCTTCATCAACCATCTCGGCGGCTTTGACCCAGTTTGGCTCTTC 370
QY 66 LeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArg 85
Db 371 CTGTGGTGGGAGAGAGATGATGTTCAATTTGGAGATTTCAGAGGTCATTCAGACGCTACGG 430
QY 86 GluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGlu 105
Db 431 GAAACACATCTTCCTCAAGTTTCTTCTGTCTGTCTCGGGAATTAATTTCTTCCTGGAG 490
QY 106 LeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValaArgAspArgPheArgGlu 125
Db 491 CTCACGTCCCGAGTTCTTACGATTTGTTTCAAAACATCGAGATCAAAAGCCAGCTGTATTTC 550
QY 126 PhePheGluSerAsnIleLysSerTyTrpArgAspAspIleAspLeuGlnAsnLeuIleAsp 145
Db 551 TTTATAAACACAAACATCAGACGATATCGGAGTACATGATTGTGCAAAACCTCATAGAC 610
QY 146 SerLeuGlnIlyrAlaAsnGlnCysGlyAlaIlyrGlyProGluAspTrpAspLeuAsn 165
Db 611 TTCACCCAGGA----- 621
QY 166 ValTyTrpPheAsnCysSerGlyAlaSerTyTrpSerArgGluIlyrCysGlyValProPheSer 185
Db 622 ATATTTCATTCAGCA-----AGTGCAGACGAGATGTGGCGATTCCATTCTCC 666
QY 186 CysCysValProAspProAlaGlnLysValValAlaAsnThrGlnCysGlyTyTrpAspValArg 205
Db 667 TGTCTGCTACTAAAGATCCCGCGAAGAGATGTCATCAACATCAAGTGTGGCTATGATGCCAGG 726
QY 206 IleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGlnAlaLeu 225
Db 727 CAAAAACCAAGAAAGTTGACCAACGACATTTGATATTACACGAAAGGCTGTGCCCCCACTTT 786
QY 226 GluSerTrpLeuProArgAsnIleTyrlIleValAlaGlyAlaPheIleAlaIleSerLeu 245
Db 787 GAGAAAGTGTTGCAGAGCAAAATTTAAACMCWGCTTGCTGTATTTTCAATAGCATTCGATTG 846
QY 246 LeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlnAlaValLys 265
Db 847 CTCGAGAAATTTGGGATTTGCTGGCCCAAGAAATTTGGTTAGCATATCGAATCGAGCTGTACG 906
QY 266 Ala 266
111

```



Db 907 GCG 909

RESULT 6

US-09-875-440-1

; Sequence 1, Application US/09875440

; Patent No. US20020156035A1

; GENERAL INFORMATION:

; APPLICANT: Reinhard, Christoph

; APPLICANT: Jefferson, Anne B.

; APPLICANT: Winter, Jill A.

; APPLICANT: Randazzo, Filippo

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: NEOPLASTIC DISEASE USING NET-4 MODULATORS

; FILE REFERENCE: PP-01701.002/200130.522

; CURRENT APPLICATION NUMBER: US/09/875.440

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 864

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: NET-4 oligonucleotide used in cell proliferation

; OTHER INFORMATION: assay on SW620 cells

US-09-875-440-1

Alignment Scores:

Pred. No.:	1.51e-59	Length:	864
Score:	567.00	Matches:	98
Percent Similarity:	77.11%	Conservative:	30
Best Local Similarity:	59.04%	Mismatches:	38
Query Match:	39.13%	Indels:	0
DB:	9	Gaps:	0

US-09-905-674-2 (1-270) x US-09-875-440-1 (1-864)

OY 6 TyrSerAsnAlaIysValSerCysTrpYrlyStryLeuLeuPheSerTyrAsnIlelle 25

DB 363 TPCAAAGGCGCCGAGGAGTGGTTCGATCAATACCTTCATATTTGGCTTCAAGTCTA 422

OY 26 PheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAlaTrpSerGluIuSgly 45

DB 423 TTTTGGTTTGGGAATACATTTCTTGGAATGTGAGCTGTGGCATGTGAAGAAAGAA 482

OY 46 ValLeuSerAspLeuThrIysValThrArgMetHisGlyIleAspProValValLeuVal 65

DB 483 GTTCGTCCCAACATCTCTTCATCCATCCGCGGCGCTTGGACCCAGTTGGCTCTTC 542

OY 66 LeuMetValGlyValValMetPheThrArgPheAlaGlyCysValGlyAlaLeuArg 85

DB 543 CTTGGTGGTGGAGGAGTGTGTTCTTTGGGATTTGGAGGGTGCATTGGAGCGCTACGG 602

OY 86 GluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGlu 105

DB 603 GAAACACATTTCTCTCAAGTTTCTGTTCTGTTCCGAGGAATATTCTTCTCTGGAG 662

OY 106 LeuAlaValAlaValLeuAlaPheLeuPheGlnAspIrrValAlaGAspArgPheArgGlu 125

DB 663 CTCACGTGCGGAGTCTAGCATTTGTTTCAAGACGTGATCAAGACACAGCTGATATTC 722

OY 126 PhePheGluSerAsnIleSeryTrpArgAspAlaPheLeuGlnAsnLeuIleAsp 145

DB 723 TTTATTAACACACATCAAGCATATGAGGATGAGCATTTGATTCGAAACCTCTATAC 782

OY 146 SerLeuGlnIysAlaAsnGlnCysCysGlyAlaIrrGlyProGluAspTrpAspLeuAsn 165

DB 783 TTCACCCAGGATATTGGAGTGCCTGTGGGCTTTTGGAGCTGATGATTTGGAACCTAAT 842

OY 166 ValTyrPheAsnCysSer 171

DB 843 ATTACTCATCATTCACACA 860

RESULT 7

US-10-052-586-277

; Sequence 277, Application US/10052586

; Patent No. US20020127584A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gunney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C1

; CURRENT APPLICATION NUMBER: US/10/052,586

; CURRENT FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059266

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/063120

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063121

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063486

; PRIOR FILING DATE: 1997-10-21

; PRIOR APPLICATION NUMBER: 60/063540

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063541

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063544

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063564

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063734

; PRIOR FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/063870

; PRIOR FILING DATE: 1997-10-31

; PRIOR APPLICATION NUMBER: 60/064103

; PRIOR FILING DATE: 1997-10-31

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066120

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/066466

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/066772

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/069335

; PRIOR FILING DATE: 1997-12-11

; PRIOR APPLICATION NUMBER: 60/069425

; PRIOR FILING DATE: 1997-12-12

; PRIOR APPLICATION NUMBER: 60/069870

; PRIOR FILING DATE: 1997-12-17

; PRIOR APPLICATION NUMBER: 60/068017

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: 60/077632

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/078886

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/078939

; PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/079664	PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786	PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107	PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080194	PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080322	PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333	PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081838	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082566	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082565	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322	PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366	PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414	PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639	PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640	PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643	PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573	PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579	PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580	PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582	PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700	PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023	PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392	PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/086609	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087599	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087608	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827	PRIOR FILING DATE: 1998-06-02

1	PRIOR FILING DATE: 1998-06-03
2	PRIOR APPLICATION NUMBER: 60/088025
3	PRIOR FILING DATE: 1998-06-04
4	PRIOR APPLICATION NUMBER: 60/088028
5	PRIOR FILING DATE: 1998-06-04
6	PRIOR APPLICATION NUMBER: 60/088029
7	PRIOR FILING DATE: 1998-06-04
8	PRIOR APPLICATION NUMBER: 60/088033
9	PRIOR FILING DATE: 1998-06-04
10	PRIOR APPLICATION NUMBER: 60/088167
11	PRIOR FILING DATE: 1998-06-05
12	PRIOR APPLICATION NUMBER: 60/088202
13	PRIOR FILING DATE: 1998-06-05
14	PRIOR APPLICATION NUMBER: 60/088212
15	PRIOR FILING DATE: 1998-06-05
16	PRIOR APPLICATION NUMBER: 60/088217
17	PRIOR FILING DATE: 1998-06-05
18	PRIOR APPLICATION NUMBER: 60/088362
19	PRIOR FILING DATE: 1998-06-04
20	PRIOR APPLICATION NUMBER: 60/088555
21	PRIOR FILING DATE: 1998-06-09
22	PRIOR APPLICATION NUMBER: 60/088722
23	PRIOR FILING DATE: 1998-06-10
24	PRIOR APPLICATION NUMBER: 60/088788
25	PRIOR FILING DATE: 1998-06-10
26	PRIOR APPLICATION NUMBER: 60/088740
27	PRIOR FILING DATE: 1998-06-10
28	PRIOR APPLICATION NUMBER: 60/088811
29	PRIOR FILING DATE: 1998-06-10
30	PRIOR APPLICATION NUMBER: 60/088824
31	PRIOR FILING DATE: 1998-06-10
32	PRIOR APPLICATION NUMBER: 60/088825
33	PRIOR FILING DATE: 1998-06-10
34	PRIOR APPLICATION NUMBER: 60/088826
35	PRIOR FILING DATE: 1998-06-10
36	PRIOR APPLICATION NUMBER: 60/088861
37	PRIOR FILING DATE: 1998-06-11
38	PRIOR APPLICATION NUMBER: 60/088863
39	PRIOR FILING DATE: 1998-06-11
40	PRIOR APPLICATION NUMBER: 60/088876
41	PRIOR FILING DATE: 1998-06-11
42	PRIOR APPLICATION NUMBER: 60/088909
43	PRIOR FILING DATE: 1998-06-12
44	PRIOR APPLICATION NUMBER: 60/089105
45	PRIOR FILING DATE: 1998-06-12
46	PRIOR APPLICATION NUMBER: 60/089512
47	PRIOR FILING DATE: 1998-06-16
48	PRIOR APPLICATION NUMBER: 60/089514
49	PRIOR FILING DATE: 1998-06-16
50	PRIOR APPLICATION NUMBER: 60/089538
51	PRIOR FILING DATE: 1998-06-17
52	PRIOR APPLICATION NUMBER: 60/089598
53	PRIOR FILING DATE: 1998-06-17
54	PRIOR APPLICATION NUMBER: 60/089653
55	PRIOR FILING DATE: 1998-06-17
56	PRIOR APPLICATION NUMBER: 60/089608
57	PRIOR FILING DATE: 1998-06-17
58	PRIOR APPLICATION NUMBER: 60/089908

Alignment Scores:

Pred. No.:	2,78e38	Length:	1778
Score:	395.50	Matches:	81
Percent Similarity:	52.55%	Conservative:	53
Best local Similarity:	31.76%	Mismatches:	114
Query Match:	27.29%	Indels:	7
DB:	12	Gaps:	3

US-09-905-674-2 (1-270) x US-10-052-586-277 (1-1778)

```

QY      14  TPTFTLYSTYRTILEUENPHESERTYASNIILEIHPHETRIPLUAIACIYVAIYALPHE 33
      |||::: ||| |||::: ::||| ||| ||| :::
Db      249  TGGGTCAAGTTTCACTTATCATCTCATTCACCGCTTTTCGGGTGATTGGGCCCTGGTCC 300
      |||::: ||| |||::: ::||| ||| ||| :::
QY      34  LEUCIYVAIGIYLEUTPALATPTPSCRIULSGIYVAIUEUSERSAPLEUTHRLYSVAI 53
      |||::: ||| |||::: ::||| ||| ||| :::

```

```

Db 309 CTGCTGTGGGCATCTATGACAGAGTTGAGCGGCGAGAAATATTAACCCCTTGAAGTGC 368
QY 54 ThrArgMetHisGlyIleAspProValIleuValIleuMetValGlyValIleuMetPhe 73
Db 369 TTC-----CTGGCTCCAGCCATCCATCCATCCCTCGGCGGTGCTCATGTTC 416
QY 74 ThrLeuGlyPheAlaGlyCysValIleAlaLeuArgIleuAsnIleCysLeuLeuAsnPhe 93
Db 417 ATGCTCTCTCTCATGTGTGTGCTGCGTCCCTCGTGACCAACCTGTACTCTCTCCAAAGC 476
QY 94 PheCysGlyThrIleValIleuLeuPhePheLeuGluIleuAlaValAlaValIleuAlaPhe 113
Db 477 TTCATGATACATCCCTGGGATCTGCCTCATCATCAGAGCTGATGTGGCTGGTGCCTTG 536
QY 114 LeuPheGlnAspTrpValAlaArgAspArgPheArgGluPhePheGluSerAsnIleLeuSer 133
Db 537 ACCCTTCGGACACACCATTCATCTCTCGAACGCAACATTCGAAAGGAGATTGAGAAC 596
QY 134 TyrArgAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnIleAlaAsnGlnCys 153
Db 597 TACTATGATGATCTGTGACTTCAAAAACATCATGACCTTGTTCAGAAAAAGTTCAGTGC 656
QY 154 CysGlyAlaIleTyrGlyProGlnAspTrpAspLeuAsnValTyrPheAsnCysSerGlyAla 173
Db 657 TGTGGCGGGGAGACTACCGAGATTGAGCAAGAAATCAGTACCAAGACTGCAGTCCCT 716
QY 174 SerTyrSerArgGluLysCysGlyValProPheSerCysValProAspProAlaGln 193
Db 717 -----GGACCCCTGGCTGTGGGTGCTGACCTCTCATCATGAGAACACGACAA 770
QY 194 LysValIleAsnThrGlnCysGlyTyrAspValArgIleGlnIleuLeuSerLysTrpAsp 213
Db 771 ---GTTGTCACACCATGTGTGGTGTACAAACTATCGACAAAGACCGTTTCAGTGTGAC 827
QY 214 GluSerIlePheThrLysGlyCysIleGlnAlaLeuGluSerTrpLeuProArgAsnIle 233
Db 828 GATGATCATCTACGTGGGGCTGCACCAACGCCGTGATCATCTGTTTCATGAGCAATAC 887
QY 234 TyrIleValIleAlaGlyValPheIleAlaIleSerIleuLeuGlnIlePheGlyIlePheLeu 253
Db 888 ACCATCATGCGCTGATCTCTCTGGGCACTCTGCTTCCCGAGTTCTGGGGGTGCTGCTG 947
QY 254 AlaArgThrIleLeuSerAspIleGluAlaValIleLysAlaGlyHis 268
Db 948 ACGCTGCTGTACATCACCCGGGTGAGAGACATCATCATGAGGACAC 992

RESULT 8
; Sequence 279, Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steve P.
; APPLICANT: Mannion, Jane
; APPLICANT: Benson, Darin R.
; APPLICANT: Carter, Darlick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864, 864
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 279
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

US-09-864-864-279
Alignment Scores:
Pred. No.: 3,16e-27 Length: 873
Score: 300.50 Matches: 72
Percent Similarity: 48.41% Conservative: 50
Best Local Similarity: 28.57% Mismatches: 97
Query Match: 20.74% Indels: 33
DB: 10 Gaps: 6

US-09-905-674-2 (1-270) x US-09-864-864-279 (1-873)
QY 16 LysTyrIleuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValIlePheLeuGly 35
Db 100 AAGTTCTTGCTCATACGCCCTCGTCCGCTTTGGCGCCGTGCATCGAGGAGACATGTGC 159
QY 36 ValGlyLeuTrpAlaTrpSerGlyLysGlyValIleuSerAspLeuThrLysValThrArg 55
Db 160 GTGGGTGTCGGGGCA-----CAGCTGTGCTGTGTCAG-----ACCATTA 198
QY 56 MetHisGlyIleAspPro-----ValIleuValIleuMetValGlyValIleVal 71
Db 199 ATCCAGGGGCTACCCCTGGCTCTCTGTCCAGTGTGCATCATCGACATGCGATGTCTTC 258
QY 72 MetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeuLeu 91
Db 259 CTCTTCCTGTGGCTTTGTGGCGCTGCTGGGGCCCGCAAGAGAACTTGTCTTATG 318
QY 92 AsnPheCysGlyThrIleValIleuIlePhePheLeuGluLeuAlaValAlaValIleu 111
Db 319 ATCAGCTTTGCCATCTTCTGTCTCTTATCATGTGTGTGAGAGTGGCGCCGACATGTGT 378
QY 112 AlaIleuLeuPheGlnAspTrpValAlaArgAspArgPheArgGluPhePheGluSerAsnIle 131
Db 379 GCGTATGTGTTAGAGTAAGTGATGTCAGAGTCTTAATTAACAATTCGCGCAGCAGATG 438
QY 132 LysSerTyrArgAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnIleAlaAsn 151
Db 439 GAGATATACCCGAAACAAACACACACGCTTCGATCTCGACAGATGACGACGATTTT 498
QY 152 GlnCysCysGlyAlaIleTyrGlyProGlnAspTrpAspLeuAsnValTyrPheAsnCysSer 171
Db 499 AAGTGTGTGGGGCTGTACTACACAGATTGGAG----- 534
QY 172 GlyIleAspTrpSerArgGluLysCysGlyValProPheSerCysValProAspPro 191
Db 535 ---AAAAATCCCTTCATGTCGAAGAACCGAGTCCCGACCTCGTGCATTAAT----- 585
QY 192 AlaGlnLysValIleAsnThrGlnCysGlyTyrAspValArgIleGlnIleuLysSerLys 211
Db 586 -----GTTACTGTGGGCTGTGGGATTAATTTCAC----- 615
QY 212 TrpAspLeuSerIlePheThrLysGlyCysIleGlnAlaLeuGluSerTrpLeuProArg 231
Db 616 ---CAGAAGCGATCCATAAGAGGGCTGTGTGAGAAATGTGGGGCTGCTGAGAGAAA 672
QY 232 AsnIleTyrIleValAlaGlyValPheIleAlaIleSerIleuLeuGlnIlePheGlyIle 251
Db 673 AATGTCTGTGTGTGATGTGAGCAGAGCCCTTGGAATGCTTTGTGACAGTTTGGGAAT 732
QY 252 PheLeuAlaArgThrIleuLeuSerAspIleGluAla 263
Db 733 GTCTTTCCTCTGCTGCTGCTGTAAGAGTATCAGAACT 768

RESULT 9
US-09-823-356-31
; Sequence 31, Application US/09823356
; Patent No. US20010025098A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.

```

```

Db      650  -----ACCTCCAGGGGTGCGGGCGCAAGCCACACCGCTTTG--- 688
Qy      210  SerLySTrpaspGluSerIlePhehTrhLySgLYCySILegLAlaLeuGluSerTrpLeu 229
Db      689  -----TGGAGACGGCGCTGCTATGAAGAGTGAAGATGTGTG 727
Qy      230  ProArgAsnIleTrIleValAlaIagLYValPheIleAlaIleSerLeuLeuGlnIlePhe 249
Db      728  GATGGCATAGACGCGTGTGGCGGGCGGTGGGATGTGCATCTCATGCATGCATCTG 787
Qy      250  gLyIlePheLeuAlaArgTrhLeuIleSerAspIle 261
Db      788  GGCATGGCTTCTTCATGACCTCTTCACAGCACATC 823

RESULT 10
US-09-783-590-10353
; Sequence 10353, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2.2
; FILE REFERENCE: PO-16,261
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420, 856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346, 731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10353
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (195)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (285)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (314)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (324)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (327)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (331)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (337)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (340)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (361)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c

```

```

: NAME/KEY: misc feature
: LOCATION: (400)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (408)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (411)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (415)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (436)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-10353

Alignment Scores:
Pred. No.: 1.53e-25 Length: 446
Score: 283.00 Matches: 60
Percent Similarity: 95.31% Conservative: 1
Best Local Similarity: 93.75% Mismatches: 1
Query Match: 19.53% Indels: 2
DB: 10 Gaps: 0

US-09-905-674-2 (1-270) x US-09-783-590-10353 (1-446)
QY 209 lysSerlylTrpAspGluSerIlePheThrlysglyCysIleGlnAlaLeuGluSerTrp 228
DB 6 AGAAGCAAGTGGAGTAGTCACATCTTCACGAAAGCGCTGACGCGCTGGAAGCTGG 65
QY 229 LeuPArg-AsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuGlnI 248
DB 66 CTCGCCGGTAACTTAACTGTGCTGGCTTCATCGGCATCTGCTTTCAGAT 125
QY 248 ePheGlyIlePheLeuAlaArgThrleuIleSerAspIleGlnAlaVal-LysAlaGlyH 268
DB 126 ATTGGCAGATCTTCCTGGCAGTACGCTGATCTCAGACATCGAGCAGTGAAGCCGGCC 185
QY 268 IShIspHe 270
DB 186 ATCACTTC 193

RESULT 11
US-09-895-828-430
: Sequence 430, Application US/09895828
: Patent No. US20020099012A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Tonglong
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Watanabe, Yoshihiro
: APPLICANT: Carter, Darrick
: APPLICANT: Henderson, Robert A.
: APPLICANT: Kalos, Michael D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.539
: CURRENT APPLICATION NUMBER: US/09/895,828
: CURRENT FILING DATE: 2001-06-28
: NUMBER OF SEQ ID NOS: 473
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 430
: LENGTH: 690
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 11, 662
: OTHER INFORMATION: n = A,T,C or G
US-09-895-828-430

Alignment Scores:
Pred. No.: 3.88e-25 Length: 690
```

```

Score: 282.00 Matches: 58
Percent Similarity: 50.25% Conservative: 43
Best Local Similarity: 28.86% Mismatches: 78
Query Match: 19.46% Indels: 22
DB: 10 Gaps: 3

US-09-905-674-2 (1-270) x US-09-895-828-430 (1-690)
QY 63 ValIleuValIleMetValGlyValAlaIlePheThrleuGlyPheAlaGlyCysValGly 82
DB 51 GTGGCATCATCGACAGTGGGTGCTTCTCTTCCTGGTGGCTTTGTGGGCTGCTGGGG 110
QY 83 AlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValIlePhe 102
DB 111 GCGTGCAAGAGAACTATTGCTTATGATCAAGTTGCCATCTTCTGCTCTTATCATG 170
QY 103 PheLeuGluLeuAlaValAlaIleuAlaPheLeuPheGlnAspTrpValArgAspArg 122
DB 171 TTGGTGAGGTGGCGCGCATTCGTGCTATGTTAGAGATAAGTGATGATGCAGAG 230
QY 123 PheArgGluPhePheGluSerAsnIleIleYserTyrArgAspAspIleAspLeuGlnAsn 142
DB 231 TTTAATAACAACCTTCGCGCAGCAGATGAGATTAACCGAAAAACAACACACTGCTTCG 290
QY 143 LeuIleAspSerLeuGlnIleAlaAsnGlnCysCysGlyValAlaTyrGlyProGluAspTrp 162
DB 291 ATCCTGGACAGAGATGACAGCATTTTAAGTCGTGGGCTGTACTACACAGATTGG 350
QY 163 AspleuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlyIleCysGlyVal 182
DB 351 GAG-----AAAATCCCTTCACATGTCGAAAGACCGAGTC 383
QY 183 ProPheSerCysValProAspProAlaGlnIleValAlaAsnThrGlnCysGlyTyr 202
DB 384 CCCGACTCTCTGTCATTAAT-----GTTACTGTGGCTGGGATT 425
QY 203 AspValArgIleGlnLeuIleYserIleTrpAspGluSerIlePheThrlysglyCysIle 222
DB 426 AATTTCAAC-----GAGAAAGCATTCATTAAGAGGGCTGTCTG 464
QY 223 GlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAla 242
DB 465 GAGAAAGATTGGGGCTGCTGAGGAAATGCTGCTGAGCTGACGACGACCCCTTGA 524
QY 243 IleSerLeuGlnIlePheGlyIlePheLeuAlaArgThrleuIleSerAspIleGlu 262
DB 525 ATTGCTTTTGTGAGGTTTGGGAATGCTTTCCTGCTGCTCGAAGAGATATCAGA 584
QY 263 Ala 263
DB 585 AGT 587

RESULT 12
US-10-042-417-37
: Sequence 37, Application US/10042417
: Patent No. US20020123082A1
: GENERAL INFORMATION:
: APPLICANT: Pagano, M.
: TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
: FILE REFERENCE: 5914-090-999
: CURRENT APPLICATION NUMBER: US/10/042,417
: CURRENT FILING DATE: 2002-01-07
: PRIOR APPLICATION NUMBER: 60/260,179
: PRIOR FILING DATE: 2001-01-05
: NUMBER OF SEQ ID NOS: 89
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 37
: LENGTH: 368
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: modified_base
```

```

; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
US-10-042-417-37

Alignment Scores:
Pred. No.:          2e-25             Length:          368
Score:              281.00            Matches:         59
Percent Similarity: 71.30%           Conservative:   18
Best Local Similarity: 54.63%        Mismatches:    28
Query Match:       19.39%            Indels:         3
DB:                12               Gaps:           0

US-09-905-674-2 (1-270) x US-10-042-417-37 (1-368)
QY      3 TTTATAGTYSERASANAALysValSeCYSTPTPLTYLSTYRLeuLeuPheSerTYR 22
        ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      49 CACGACGATTTCACGAACAAGCTGAAGTGGCGTCTGGCGGAAATCACTTCGTGGCCTTC 108
        ||||| : | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY     23 ASNllellepHeTRPeuAlAglYalVAlpheleuSlYalJAlyleuTPAlaTRPSer 42
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     109 AACATTCTTCCTTGCGTGGCTGGAGCCCTGTCTGCCTGCTATCGGCCCTCGGCGTGGG 168
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY     43 GLUylSGlYAlleuserrAsPlenThrLYsVALPhARgmethISgLYleasPRovAl 62
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     169 GAGAAGGCGCTCTCTCGAACATCTCACGCCCTGACAgATtGGAGAGCCTTGACCOCGTg 228
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY     63 VALleuVal-LeumETyAlGlYAlYAlValmetPheThrLeuGlYPheaLA-GlyCS-Val 81
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     229 TGCGTTGTTGTGTGTACTGTGGAGCGCATGTCGGTCTGCGCTGTTCGTGGCTGCATT 288
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY     82 GLYAlAleuArGGLuASNlleCYSLeuLeuAsnPhePhecSclYthrIleValleulle 101
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     289 GGGGCGCTCGGGGAACACCTTCCTGCTCAAGTTTTCTNCNGNTTCTCGGTCTCATC 348
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY     102 PhePheLuengluLeuAlA 107
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     349 TTCTTCCTCGAGCTGGCA 366
        |||||:::| | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-09-729-674-89
; Sequence 89, Application US/09729674
; Patent No. US2001003935A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: LaValle, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Reichel, Kim
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6055-64X
CURRENT APPLICATION NUMBER: US/09/729,674
CURRENT FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: 09/539,330
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 283
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 89
LENGTH: 1797
TYPE: DNA
ORGANISM: Homo sapiens
US-09-729-674-89

Alignment Scores:
Pred. No.:          1.75e-23             Length:          1797
Score:              1.75e-23            Matches:         59
Percent Similarity: 71.30%           Conservative:   18
Best Local Similarity: 54.63%        Mismatches:    28
Query Match:       19.39%            Indels:         3
DB:                12               Gaps:           0
```

Percent Similarity:	273.50	Matches:	67
Score:	44.49%	Conservative:	46
Best Local Similarity:	26.38%	Mismatches:	100
Query Match:	18.8%	Indels:	41
DB:	10	Gaps:	8

  

US-09-905-674-2 (1-270) x US-09-729-674-89 (1-1797)	
Oy	11 ValsercysrtprrylrlystlyrleuLeuPheSerlyrnslllellePheTrpleuAlaGly 30
Db	54 ATTCCTT---TTCAAGAGCGTCTCTCAATCTACACTTTTATTTTCTGGATCACTGGC 110
Oy	31 ValValPheLeuGlyValGlyLeuTrpAlaTrpSerGlyluysGlyValLeuSerAspLeu 50
Db	111 GTTATTCCTTTCGGAGTTGGCATTTGGGGC-----AAGGTAGACCTTGAGAAATTAC 161
Oy	51 ThrlyValIlnrArGmethIsGlyIleAspProValValLeuValLeuMetValGlyVal 70
Db	162 TTTTCTCTTTTAAATGAGAAAGCCACCAAACTGCCCTTGTCCTCATCTGCTACTGGTACC 221
Oy	71 ValMetPheTrpLeuGlyPheAlaGlyCysValGlyAlaLeuArGluAsnIleCysLeu 90
Db	222 GTCATATTCTTTTGGGACCTTTGGTTGTTTCTACCTGCCGAGCTTCTGCATGGATG 281
Oy	91 LeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaVal 110
Db	282 CTAAAAACGTATGCAGATCTTCTACCTCTGTTTGGTTCGCAACTGGTCGCTGCATC 341
Oy	111 LeuAlaPheLeuPheGlnAsnPrrValArGAsprrPheArGluPhePheGlnSerAsn 130
Db	342 GTAGCATTTGTTTTCAGACATGACATTAAAGACGCTTTAAAGAAATTAATTGAGAAAGCT 401
Oy	131 IleuSerTyrrArGAspAspIleAspLeuGln---AsnLeuIleAspSerLeuGlnLys 149
Db	402 TTGAAGCAGATTAATCTACAGGAGATTATAGAAAGCCATGCAGTAGCAAGATCCAAAT 461
Oy	150 AlaAsnGlnCysGlyAlaTyrglyProGluAspTrpAspLeuAsnValTyrrPheAsn 169
Db	462 ACGTTCATTTTGTGGTCCACCAGATTATAGACATTGG----- 500
Oy	170 CysSerGlyAlaSerTyrrSerArGluLysCysGlyAlaProPheSerCys----- 186
Db	501 ---ACAGTACTATTATTACTCGAAGAAA---CGATTCTTAAGAGTTCTCTTAACCT 554
Oy	187 ----CysValProAsp--ProAlaGlnLysValAlaAsnThrGlnCysGlyTyrrAsp 203
Db	555 GAAGATTCTACTCCACAGAGAGATGACAGCAAAATAAACAAT----- 596
Oy	204 ValArGlnIleGlnLeuLysSerlyTrpAspGluSerIlePheThrlysglyCysIleGln 223
Db	597 -----GAGGTGTTTATTATA 611
Oy	224 AlaLeuGlnSerTrpLeuProArGAsnIleTyrrIleValAlaGlyAlaPheIleAlaIle 243
Db	612 AAGGTAGTACCATTATAGACTGAGAATGGAGACTCGTTCAGAAATTCTCTTGAGATT 671
Oy	244 SerLeuLeuGlnIlePheGlyIlePheLeuAlaArGhrLeu 257
Db	672 GCTTGCTTCCAACTGATTGAATCTTCTGCTACCTGCTC 713

  

RESULT 14	
US-09-823-356-32	
; Sequence 32. Application US/09823356	
; Patent No. US20010025098A1	
; GENERAL INFORMATION:	
; APPLICANT: Tang, Y. Tom	
; APPLICANT: Bandman, Olga	
; APPLICANT: Lai, Preeti	
; APPLICANT: Hillman, Jennifer L.	
; APPLICANT: Yue, Henry	
; APPLICANT: Corley, Neil C.	
; APPLICANT: Guegler, Karl J.	
; APPLICANT: Kaser, Matthew R.	

```

; APPLICANT: Baughn, Mariah R.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN MEMBRANE SPANNING PROTEINS
; FILE REFERENCE: PF-0489-1 CON
; CURRENT APPLICATION NUMBER: US/09/823,356
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/039,307
; PRIOR FILING DATE: 1998 March 13
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PERL Program
; SEQ ID NO: 32
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID NO. US20010025098A1 2779610
US-09-823-356-32

```

```

Alignment Scores:
Pred. No.: 2,63e-23 Length: 2358
Score: 273.50 Matches: 67
Percent Similarity: 44.48% Conservative: 46
Best Local Similarity: 26.38% Mismatches: 100
Query Match: 18.88% Indels: 41
DB: 10 Gaps: 8

```

US-09-905-674-2 (1-270) x US-09-823-356-32 (1-2358)

```

QY 11 ValSerCysTrpTyrIleuLeuPheSerTyrAsnIleIlePheThrLeuAlaGly 30
DB 156 ATTACTTGT--TTCAGAGCGCTTCTGTAATCTACACTTTATTTCTGATCGATGCG 212
QY 31 ValValPheLeuGlyValGlyLeuTrpAlaTrpSerGluValGlyValLeuSerSpleu 50
DB 213 GTTATCTCTTTCGACGTTGGCAATTGGGCG-----AAGGTGAGCGCTGGAGAAATTAC 263
QY 51 ThrIysValThrArgMetHisGlyIleAspProValIleuValIleuMetValGlyAl 70
DB 264 TTTTCCTTTTAAAGAGAGAGCCACCAATGCCCCCTCGTCTGCTACTGCTGCTAC 323
QY 71 ValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeu 90
DB 324 GCATATTATCTTTTGGCGACCTTGTGTTGCTTACCTGCCGAGCTTTCGATGATG 383
QY 91 LeuAsnPhePheCysGlyThrIleValIleuIlePhePheLeuGluAlaValAlaVal 110
DB 384 CTAAACCTGTATGCAATGTTTCTGACTCGTTTGTGTCGACGCTGCTGCTGCTG 443
QY 111 LeuAlaPheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGluSerAsn 130
DB 444 GTAGATTGTTTTCAGACATGAGATTAAAGAACCTTTAAGATTAATTATGAGAGGCT 503
QY 131 IleIysSerTyrArgAspSpleuGln--AsnLeuIleAspSerLeuGlnLys 149
DB 504 TTGAAGACGTATACCTACAGAGAGATTATAGAACCATGACGTAGACAAAGTCAAAAT 563
QY 150 AlaAsnGlnCysCysGlyValAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsn 169
DB 564 ACCTTGCTGTTGTTGGTGTCCAGCATTAAGAGATTG----- 602
QY 170 CysSerGlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCys----- 186
DB 603 ---ACAGATACATAATTATTAACAGAAAA---GGATTTCCTAAGAGTTGCTGTAACCT 656
QY 187 -----CysValProAsp---ProAlaGlnLysValIleAsnThrGlnCysGlyTyrAsp 203
DB 657 GAAGATTGTACTCCAGAGAGATGACAGAAAGTAACAAAT----- 698
QY 204 ValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGln 223
DB 699 -----GAAGTTGTTTAAVA 713

```

```

QY 224 AlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAlaIle 243
DB 714 AAGGTGATGACCATTTATAGACTCAGAAATGGGAGCTGTTGAGGAATTTCTTGGAGTT 773
QY 244 SerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeu 257
DB 774 GCTTGCTTCCAACTGATGTAATCTTCTGCGCTACAGCCGC 815

```

#### RESULT 15

```

US-09-925-301-392
; Sequence 392, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 392
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (25)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (54)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (38)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-392

```

#### Alignment Scores:

```

Pred. No.: 1.13e-22 Length: 1545
Score: 266.00 Matches: 74
Percent Similarity: 44.48% Conservative: 51
Best Local Similarity: 26.33% Mismatches: 108
Query Match: 18.36% Indels: 48
DB: 10 Gaps: 11

```

US-09-905-674-2 (1-270) x US-09-925-301-392 (1-1545)

```

QY 6 TyrSerAsnAlaLysValSerCys-----TrpTyrLysTyrLeuLeuPheSerTyr 22
DB 98 TTCAAGACGAGAAAGAAACATGTGGACCGCTTGGCTCAAGTACCTGTTTACTTAC 157
QY 23 AsnIleIlePheThrPheLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAlaTrpSer 42
DB 158 AATTGCTGCTTCTGCGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCC 217
QY 43 GluLysGly---ValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspPro 61
DB 218 CTCAAGAGTACACTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 268
QY 62 ValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysVal 81
DB 269 GCCTACATCTCGTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 328
QY 82 GlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValIleGln 101
DB 329 GCCACCTTCAAGAGCGCTCGGAACCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 388

```

```
QY 102 phepheleugluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTyrValArgAsp 121
   ||| |||||::: ::|||::: :::
Db 389 TTTCGTGCGAGATCATCGCTGCTATCCTCGCTACGCCCTACTACAGCAGCTGAACAG 448
QY 122 ArgPheArgGluPhePheGlnSerAsnIle--LysSerTyrArgAsp-----AspIle 138
   ::||| ::: ||| |||
Db 449 GAGCTCAAGGAGAACCTGAGGACACCATGACCAAGCGCTACACCAAGCCGGCCATGAG 508
QY 139 AspleuglnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaIleTyrGly 158
   ||| ::: ||| |||||::: |||||:::
Db 509 GCTGTGACCAAGCGCTGTGACCAAGCTGCAAGAGTTCCACTGTGTGCGACACACAC 568
QY 159 ProGlnAspTyrPaspLeuAsnValTyrPheAsnGlnCysSerGlyAlaSerTyrSerArgGlu 178
   ::|||::: ::: ::: |||||:::
Db 569 TCACAGAGACTGCGCGAGACAGATGATGATCCGC-----TCACAGAG 610
QY 179 LysCysGly-----ValProPheSerCysValProAspProAlaGlnLysValVal 196
   ||| ||| ||||| |||||
Db 611 GCCGGTGGCCGTGTGTGTCACAGACAGCTGCTG-----AAGACGGTG 652
QY 197 AsnThrGlnCysGly-----TyrAspValArgIleGlnLeu 208
   ||||| ||| |||
Db 653 GTGGCTCTTTGTGGACAGGAGACCATGCCCTCCACACATCTACAGGTGAG----- 703
QY 209 LysSerLysTyrPaspGlnSerIlePheThrLysGlyCysIleGlnAlaLeuGlnSerTyr 228
   ||||| ||||| |||||
Db 704 -----GGCGGCTGCATCACCAGTTGGAGACCTTC 733
QY 229 LeuProAlaGlnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIle 248
   ::||| ::|||::: ||| ||| |||||:::
Db 734 ATCCAGAGACACCTGAGGGCTGATGGGGCTGCGGATCGGCATGCTGCTGTGCAGGCTC 793
QY 249 PheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlnAlaValLysAlaGlyHis 268
   |||||::: ||| ||| ::|||::: |||
Db 794 TTGGCATATCTTCACGTGCTGCTG-----TACAGGAGTCTCAAGCTGGAGCAC 844
QY 269 His 269
   ::|
Db 845 TAC 847
```

Search completed: December 19, 2002, 08:43:59  
Job time : 66 secs

. . . .